

LITHOLOGICAL CLASSIFICATION BY DEEP LEARNING ALGORITHM



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Earth Sciences

Department of Geology

Faculty of Science

Chulalongkorn University

Academic Year 2018

Copyright of Chulalongkorn University

การจำแนกวิทยานิพนธ์ด้วยขั้นตอนวิธีการเรียนรู้เชิงลึก



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาโลกศาสตร์ ภาควิชาธรณีวิทยา
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2561
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บรรณ ทงเสม : การจำแนกวิทยาหินด้วยขั้นตอนวิธีการเรียนรู้เชิงลึก . (LITHOLOGICAL CLASSIFICATION BY DEEP LEARNING ALGORITHM) อ.ที่ปรึกษาหลัก : ผศ. ดร.วรัญทร คณิตปัญญาเจริญ, อ.ที่ปรึกษาร่วม : อ. ดร.เอกพล ช่วงสุวนิช

งานวิจัยฉบับนี้ได้นำขั้นตอนวิธีการเรียนรู้ของเครื่อง 6 วิธี มาใช้ในการแบ่งชนิดหินจากข้อมูลการหยั่งธรณีหลุมเจาะ 15 ประเภท จาก 4 หลุมเจาะ ในบริเวณที่ราบแม่น้ำสเนก รัฐโอดาโฮ ประเทศสหรัฐอเมริกา สำหรับการทดลองได้แบ่งออกเป็น 2 แบบ คือ การทดสอบแบบหลุมเดี่ยวและการทดสอบแบบหลายหลุม การทดสอบแบบหลุมเดี่ยวจะนำข้อมูลของแต่ละหลุมมาแบ่งเป็น ข้อมูลชุดอบรม (ร้อยละ 70) ข้อมูลชุดตรวจสอบ (ร้อยละ 10) และข้อมูลชุดทดสอบ (ร้อยละ 20) การทดสอบแบบหลายหลุมจะรวมข้อมูลจาก 3 หลุมแรกและร้อยละ 70 ของข้อมูลหลุมที่ 4 เป็นข้อมูลชุดอบรม และข้อมูลส่วนที่เหลือจะเป็นข้อมูลชุดตรวจสอบและทดสอบต่อไปในร้อยละ 10 และ 20 จากการทดลองพบว่าแบบจำลองเอ็กซ์ตรีมเกรเดียนต์บูตติ้งให้ค่าความถูกต้องสูงที่สุดคิดเป็นร้อยละ 91 ในแบบหลุมเดี่ยว และร้อยละ 87 ในแบบหลายหลุม เนื่องจากแบบจำลองนี้สามารถเลือกใช้ข้อมูลที่มีประโยชน์ในการจำแนกชนิดหิน และจัดการกับข้อมูลที่ขาดหายไปได้ด้วยการตัดสินใจแบบโครงสร้างต้นไม้ นอกจากนี้งานวิจัยชิ้นนี้ยังแสดงให้เห็นว่าการทดสอบแบบหลายหลุมมีความซับซ้อนมากกว่าการทดสอบแบบหลุมเดี่ยว เนื่องจากการรวมข้อมูลจากหลายหลุมเป็นการรวมข้อมูลหลายรูปแบบ ซึ่งส่งผลต่อการตัดสินใจของแบบจำลองแบบจำลองโครงข่ายประสาทเทียมซึ่งเป็นแบบจำลองการเรียนรู้เชิงลึกชนิดหนึ่ง ให้ค่าความถูกต้องที่ต่ำกว่าเอ็กซ์ตรีมเกรเดียนต์บูตติ้งทั้ง 2 การทดลอง เนื่องจากแบบจำลองชนิดนี้ไม่สามารถจัดการกับข้อมูลที่ ไม่สมดุลได้ดีเท่ากับเอ็กซ์ตรีมเกรเดียนต์บูตติ้ง ในภาพรวมทุกแบบจำลองสามารถแยกแยะหินอัคนีด้วยค่าความถูกต้องที่สูง เนื่องจากปริมาณของข้อมูลที่มีจำนวนมาก ทำให้แบบจำลองเรียนรู้ลักษณะเฉพาะของหินอัคนีได้อย่างมีประสิทธิภาพ หินตะกอนซึ่งมีปริมาณข้อมูลน้อย (ร้อยละ 6) และค่าของข้อมูลหยั่งธรณีหลุมเจาะที่ใกล้เคียงกันมากทำให้การแยกแยะชนิดหินเป็นไปได้ยาก อย่างไรก็ตามการแยกแยะหินตะกอนด้วยขั้นตอนวิธีการเรียนรู้ของเครื่องสามารถปรับปรุงให้ดีขึ้นได้ โดยการเพิ่มจำนวนข้อมูลหินตะกอนและการประยุกต์ใช้ข้อมูลชนิดอื่นเพิ่มเติม เช่น ข้อมูลขนาดตะกอน เป็นต้น

สาขาวิชา	โลกศาสตร์	ลายมือชื่อนิสิต
ปีการศึกษา	2561	ลายมือชื่อ อ.ที่ปรึกษาหลัก
		ลายมือชื่อ อ.ที่ปรึกษาร่วม

6072177023 : MAJOR EARTH SCIENCES

KEYWORD: Well logging, Snake River Plain, Deep learning, Machine Learning

Worapop Thongsame :
LITHOLOGICAL CLASSIFICATION BY DEEP LEARNING ALGORITHM. Advisor: Asst.
Prof. Waruntorn Kantipanyacharoen, Ph.D. Co-advisor: Ekapol Chuangsuwanich,
Ph.D.

Six machine learning algorithms are used to classify subsurface rocks based on fifteen well logging features from four geothermal wells in the Snake River Plain, Idaho, USA. Two experimental designs, single- and multiple-well tests, are developed to determine the most optimal model and hyperparameters. The single-well test randomly assigns the data in each well into 70% for training set, 10% for validation set, and 20% for test set. The multiple-well test combines data from three wells and splits the data in the fourth well into 70% for training set, 10% for validation set, and 20% for test set. Results show that Extreme gradient boosting model (XGB) gives the highest accuracies in single- and multiple-well tests at 91% and 87%, respectively. This is because XGB can avoid unnecessary features and missing values based on decision tree classifier. In addition, multiple-well test is more complex and generally gives lower prediction accuracy than those of single-well test due to the variety of features from different wells. Artificial neural network (ANN), one of the deep learning algorithms, consistently gives lower accuracy than that of XGB in both tests. This is because ANN cannot handle imbalanced dataset as well as XGB. Overall, igneous rocks can be accurately classified due to their abundance, which allows the models to effectively learn about their distinct characteristics. Sedimentary rocks are the minor classes and mostly contain overlapped well logging responses, which impose difficulty in lithological classification. The classification of sedimentary rocks can be further improved by increasing a number of data and incorporating other physical properties such as grain size.

Field of Study: Earth Sciences Student's Signature

Academic Year: 2018 Advisor's Signature

Co-advisor's Signature

ACKNOWLEDGEMENTS

There is much assistance for me during the last two years to do this research. First, the completion of this work could not have been possible without the well logging data of the Snake River Plain published by Idaho National Laboratory. I would like to thank my advisor Assistance Professor Waruntorn Kanitpanyacharoen, Ph.D. of the Department of Geology, Faculty of Science, Chulalongkorn University. The door of her office was always open for me whenever I faced with a hard question about research or my daily life. I would also like to thank the expert of data science who has been involved in my thesis for two years: my co-advisor Ekapol Chuangsuwanich, Ph.D. Without his professional participation and suggestion, this thesis could not have successfully conducted. Last but not least, it is a great pleasure to thank my idol who has provided unending inspiration and encouragement: Cherprang Areekul, BNK48. She taught me that every effort will always be rewarded. Thank to her statement, I have tried again and again until this thesis was complete. Without her inspiration and encouragement, this thesis would not be complete and I would not be here.

Worapop Thongsame

TABLE OF CONTENTS

	Page
.....	iii
ABSTRACT (THAI).....	iii
.....	iv
ABSTRACT (ENGLISH).....	iv
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
List of tables.....	x
List of figure.....	xi
Chapter 1 Introduction.....	1
1.1 Introduction.....	1
1.2 Objectives.....	4
1.3 Scope of works.....	4
1.4 Expected results.....	5
Chapter 2 Literature Reviews.....	6
2.1 Well logging.....	6
2.1.1 Resistivity log.....	7
2.1.2 Gamma ray log.....	8
2.1.3 Acoustic log.....	9
2.1.4 Density log.....	10
2.1.5 Neutron log.....	12
2.1.6 Direct measurement log.....	12

2.2 Machine learning algorithms	13
2.2.1 Previous works	13
2.2.2 Support vector machine (SVM)	14
2.2.3 K nearest neighbour (K-nn).....	16
2.2.4 Extreme gradient boosting (XGB).....	18
2.2.5 Artificial neural network (ANN)	20
2.2.6 Convolutional neural network (CNN).....	22
2.2.7 Gated recurrent unit (GRU).....	24
2.3 Experimental design	25
Chapter 3 Study Area	26
3.1 Overview of data.....	26
3.2 Snake River Plain.....	35
3.2.1 Western Snake River Plain.....	38
3.2.2 The Eastern Snake River Plain.....	44
Chapter 4 Methodology.....	48
4.1 Overview of Workflow.....	48
4.2 Data preparation	51
4.3 Models development	53
4.3.1 Experimental design	53
4.3.2 Hyperparameter tuning.....	55
4.4 Models comparison.....	57
Chapter 5 Results	58
5.1 Model Performance.....	58
5.1.1 Single-well test	59

5.1.2 Multiple-well test.....	61
5.2 The effect of tuning parameter	63
5.3 Confusion matrix	67
Chapter 6 Discussion.....	72
6.1 Model performance.....	72
6.1.1 Single-well test	72
6.1.2 Multiple-well test.....	74
6.2 The effects of tuning parameters and amount of data	78
6.3 Feature importance and ranges	83
Chapter 7 Conclusions	88
7.1 Conclusions	88
7.2 Recommendations.....	89
APPENDIX A.....	91
APPENDIX B	92
APPENDIX C	93
APPENDIX D.....	94
APPENDIX E	95
APPENDIX F	96
APPENDIX G.....	97
APPENDIX H.....	98
APPENDIX I.....	99
APPENDIX J.....	100
APPENDIX K.....	101
REFERENCES	102

VITA..... 111



จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

List of tables

	Page
Table 1 Available well-logs (features) in WO2, Mountain Home (MH), Kimma, Kimberly in the Snake River Plain.	28
Table 2 Rock types (classes) in WO2, Mountain Home (MH), Kimma, Kimberly in the Snake River Plain.....	29
Table 3 A range of hyperparameter tuned for each algorithm.....	56
Table 4 Classification accuracies of each algorithm for single- and multiple-well tests.	59
Table 5 The representative accuracy of each rock type from the single-well test. The representative accuracy is averaged from all wells and selected from one out of five folds in Monte Carlo cross-validation.....	60
Table 6 The representative accuracy of each rock type from the multiple-well test. The representative accuracy is averaged from all wells and selected from one out of five folds in Monte Carlo cross-validation.....	62
Table 7 The optimal range of hyperparameter tuned for each algorithm in single- and multiple-well tests.....	64
Table 8 Range of well logging values in each rock from WO2 well in single-well test by XGB compared to the other works.	84
Table 9 Range of well logging values in each rock from WO2 well in single-well test by XGB compared to the other works (continued).....	85

List of figure

	Page
Figure 1 An example of well logging data of volcanic rocks in the Anda Sag, the Songliao Basin (Huang et al., 2015).	7
Figure 2 An example of gamma ray log correlated with the sedimentary rock section in Loranca Basin, Spain (Martinius et al., 2002).	9
Figure 3 The well logging data, gamma ray log (solid green line), P-wave sonic log (solid red line), and S-wave sonic log (solid blue line) of the upper and lower Marcellus shale and Cherry Valley limestone (Roy, 2013).	10
Figure 4 A correlation of well logging data, gamma ray log, density log, and neutron log, with lithology in the Rift System, Kansas (Berendsen et al., 1988).....	11
Figure 5 A case in point of SVM with soft margin. Right represents data set which has noisy data. Left shows separating hyperplane divided classes with having some misclassified data points.....	15
Figure 6 The concept of SVM with RBF. Left is SVM with soft margin and black line is decision boundary between Class 1 and Class 2 . Right is SVM with RBF. Pink and green surface are decision boundary between Class 1 and Class 2.	16
Figure 7 Classification of unlabeled data point by majority vote. The unlabeled is classified to Class 1 because Class 1 is the most popular class among the K nearest data points.	17
Figure 8 K nearest neighbour with too small K (a), suitable K (b), and too large K (c).17	
Figure 9 The procedures of decision tree. First, training data are used to construct a decision tree and then it is applied to classify new data point from test data.....	18
Figure 10 The concept of overfitting and underfitting. On the left, model cause underfitting so decision boundary	20

Figure 11 A analogy between nerve system (top) and Artificial neural network (bottom).....	21
Figure 12 Dropout parameter. Dropout eliminates some of nodes while training step to prevent model to over-reliance on some nodes.....	22
Figure 13 The structure of CNN. It consists of input layer, convolutional layer, pooling layer, and fully connected layer.....	23
Figure 14 Convolutional operation.....	24
Figure 15 The conceptual model for GRU.....	25
Figure 16 Location of 4 study wells. Yellow spots represent the location of study well. WO2, Kimma, and Kimberly well stand on Eastern Snake River Plain while Mountain Home locate on Western Snake River Plain (Shervais et al., 2013).....	26
Figure 17 Classes distribution of each well. BS: basalt, CS: claystone, CG: conglomerate, SS: sandstone, ST: siltstone, TF: tuff, VP: vitrophyre, SR: sedimentary rock, and RH: rhyolite.	30
Figure 18 Well logging data of WO2 with lithostratigraphic log. This well depth is about 5000 m with 0.5 m interval. GR: gamma ray, DENS: density, POR: porosity, and NEUT: neutron.....	31
Figure 19 Well logging data of Mountain Home with lithostratigraphic log. GR: gamma ray, Temp: temperature, P: pressure, Rmud: resistivity in mud measurement, Rd: resistivity in deep measurement, Rs: resistivity in shallow measurement, Th: thorium, U: Uranium, K: Potassium, Vp: P-wave velocity, Vs: S-wave velocity, Vw: water wave velocity.	32
Figure 20 Well logging data of Kimma with lithostratigraphic log. There are about 800 m in Kimma well depth with 0.1 m interval. GR: gamma ray, Temp: temperature, P: pressure, Rd: resistivity in deep measurement, Rs: resistivity in shallow measurement, Vp: P-wave velocity, Vs: S-wave velocity.....	33
Figure 21 Well logging data of Kimberly with lithostratigraphic log. Kimberly well is drilled approximately 1,900 m with 0.1 m interval. GR: gamma ray, Temp:	

temperature, P: pressure, Rmud: resistivity in mud measurement, Rd: resistivity in deep measurement, Rs: resistivity in shallow measurement, Th: thorium, U: Uranium, K: Potassium, Vp: P-wave velocity, Vs: S-wave velocity, Vw: water wave velocity.....	34
Figure 22 Crustal thickness map of western United State showing tectonic features and physiographic province: Basin and Range (B&R), Blue Mountains (BM), Cascade Arc (CA), Coast Ranges (CR), Colorado Plateau (CP), Columbia Plateau (CoP), Great Plains (GP), Great Valley (GV), High Lava Plains (HLP), Klamath Mountains (KM), North Cascades (NC), Northern Rocky Mountains (NRM), Olympic Mountains (OM), Sierra Nevadas (SN), Snake River Plain (SRP), Southern Rocky Mountains (SRM), and Wyoming Basin(WB) (Bedrosian and Feucht, 2014).	36
Figure 23 A geological map of Idaho, the USA, Snake River Plain with well location. This area is covered by Quaternary basalt and rhyolite (Lewis et al., 2012).	37
Figure 24 The Snake River Plain in south Idaho, USA. Red dash line represents the boundary between the Eastern Snake River Plain and the Western Snake River Plain. Black dash line refers to the area of Quaternary volcanic activity (Neupane et al., 2014).	38
Figure 25 The intracontinental rift basin Western Snake River Plain. Black lines are border normal fault dipping toward the center of Western Snake River Plain (Shervais and Vetter, 2009).	39
Figure 26 Cross section of Western Snake River Plain. The dark pattern shows basalt in this area (Wood and Clemens, 2002).	40
Figure 27 Stratigraphic column of Chalk Hills Formation represented rock in the Western Snake River Plain (Wood and Clemens, 2002).	43
Figure 28 Eastern Snake River Plain bounded by Basin and Range and seismic parabola (green line) (Rodgers et al., 2002).	44
Figure 29 The typical cross section of Eastern Snake River Plain. There are more volcanic eruptions in this area than Western Snake River Plain.	45

Figure 30 Evolution of Eastern Snake River Plain in stage 1 to 4 proposed by Rodgers et al. (2002).....	47
Figure 31 Data preparation step. The data are cleaned and split into training, validation, and test data.....	48
Figure 32 Models development step. The models are trained by training data and optimized by validation data. After that, they are evaluated by test data.....	49
Figure 33 Models comparison step. The performance of six models which are optimized by validation data is compared in this step.....	50
Figure 34 The conceptual model of Class Center based Missing Value Imputation. It calculates missing values for surrounding data points similar to K-nn (Tsai et al., 2018).....	52
Figure 35 A diagram shows stratified sampling where the data are split while class distribution remains constant.....	52
Figure 36 The single-well test. The training (70%), validation (10%), and test (20%) data are randomly split from the data in each well.	54
Figure 37 The multiple-well test. Three wells and 70% of the fourth well are assigned as training data while the rest of the fourth well is assigned as validation (10%) and test (20%) data.....	55
Figure 38 The comparison results between well report and the models. XGB gives the highest accuracy score at 90%. SVM and K-nn provide comparable results at 89% while ANN, CNN, and GRU exhibit the lowest accuracy.....	58
Figure 39 Average accuracy scores over six machine learning algorithms in single-well test. Red spots represent min and max accuracies of each algorithm before average. Error bars represent standard deviation.....	60
Figure 40 Average Accuracy scores over six machine learning algorithms in multiple-well test. Red spots represent min and max accuracies of each algorithm before average. Error bars represent standard deviation.	62

Figure 41 The line graphs show the effect of tuning parameter to accuracy in XGB in two experiments. Red line represents the accuracy of test data and validation data for green line.	66
Figure 42 Normalized confusion matrix of six models in prediction of WO2 well in the single-well test. BS: basalt, CS: claystone, CG: conglomerate, SS: siltstone, ST: sandstone, TF: tuff, VP: vitrophyre.	68
Figure 43 Normalized confusion matrix of six models in prediction of Kimberly well in the single-well test. BS: basalt, SR: sedimentary rock, RH: rhyolite.	71
Figure 44 Comparison of accurate performance between this study and other studies in single-well test. Red color represents this study. Black refers to a study by Xie et al. (2018). Green and blue show results from Dubois et al. (2007) and Konaté et al. (2015), respectively.	74
Figure 45 Comparison of accurate performance in multiple-well test between this study and Hall and Hall (2017).	75
Figure 46 Comparison average accuracy of each model in this study between single-well test and multiple-well test.	75
Figure 47 Comparison accuracy of each rock type in Kimberly well between single well and multiple-well test in this study. The accuracy is provided by XGB.	76
Figure 48 The used features by XGB in the single-well test. Feature score represents the importance of each feature in classification.	78
Figure 49 The effect of tuning parameter to the accuracy of XGB in single-well test. Red lines represent test subset and green lines represent validation subset.	79
Figure 50 A comparison of accuracy between ANN and XGB with a number of data represented by size of spot. ANN cannot classify minor classes correctly while XGB can.	82
Figure 51 Well logging marks of volcanic rocks in the Anda Sag, the Songliao Basin (Huang et al., 2015).	86

Figure 52 Average feature importance scores of each feature or type of well logs in XGB algorithm from multiple-well test.	87
Figure 53 The effect of tuning parameters in SVM.	91
Figure 54 The effect of tuning parameter in K-nn.	91
Figure 55 The effect of tuning parameters in ANN.	92
Figure 56 The effect of tuning parameters in CNN.	93
Figure 57 The effect of tuning parameters in GRU.	94
Figure 58 Normalized confusion matrix of six models in prediction of Mountain Home well in the single-well test. BS: basalt, SR: sedimentary rocks.	95
Figure 59 Normalized confusion matrix of six models in prediction of Kimma well in the single-well test. BS: basalt, SR: sedimentary rocks.	96
Figure 60 Normalized confusion matrix of six models in prediction of WO2 well in the multiple-well test. BS: basalt, SR: sedimentary rocks, VP: vitrophyre, TF: tuff, RH: rhyolite.	97
Figure 61 Normalized confusion matrix of six models in prediction of Mountain Home well in the multiple-well test. BS: basalt, SR: sedimentary rocks, VP: vitrophyre, TF: tuff, RH: rhyolite.	98
Figure 62 Normalized confusion matrix of six models in prediction of Kimma well in the multiple-well test. BS: basalt, SR: sedimentary rocks, VP: vitrophyre, TF: tuff, RH: rhyolite.	99
Figure 63 Normalized confusion matrix of six models in prediction of Kimberly well in the multiple-well test. BS: basalt, SR: sedimentary rocks, VP: vitrophyre, TF: tuff, RH: rhyolite.	100
Figure 64 An example of well logging data in excel file from Mountain Home.	101

Chapter 1 Introduction

1.1 Introduction

Subsurface information is essential data for understanding geological evolution. Lithological information is used to explore various natural resources such as groundwater, petroleum, and geothermal energy. Collecting rock samples by drilling is a direct method to obtain subsurface information but this process is expensive and time-consuming. Geophysical surveys such as resistivity, gravity, and well logging exploration are invented to acquire subsurface information. In particular, well logging exploration has been widely used to gain physical properties of subsurface rocks. Main advantages of well-logging include highly vertical resolution, good continuity, and convenient data acquisition (Xie et al., 2018). Compared to other geophysical methods such as resistivity, gravity, and magnetic surveys, well logging has a high vertical resolution (Soltani et al., 2016). The data is usually collected at every 30 - 50 cm. Well-logging data is also continuous and can be collected while drilling and/or after drilling. However, there are no logging tools which can directly determine physical properties of rock, for example, porosity, permeability, and lithology.

Various types of logging tool are used to collect geophysical signals of the subsurface formation (Xie et al., 2018). For instance, the gamma ray log detects an amount of gamma ray emitted from radioactive elements in the rock formation to infer volume of them. Resistivity log measures the electrical resistance of rock formation in order to investigate types of pore fluid. Temperature log detects the temperature of the rock formation to calculate heat transfer. After a massive amount of subsurface information is gathered, it is interpreted to gain insights into the lithology and physical properties of the rock formation. However, the criteria of lithological classification are loosely defined and human error can significantly

contribute to the uncertainty of the interpretation. Previously there have been many mathematical methods such as deconvolution, noise filter, and signal stacking applied to well logging interpretation to reduce the error and enhance well logging signal (Dubois et al., 2007). Recently, machine learning algorithms have been introduced to solve problems in pervasive fields, for example, disease detection, climate modeling, and also well logging interpretation (Catalogna et al., 2012; Devak et al., 2015; López et al., 2017; Wang et al., 2013). Machine learning algorithms use statistical techniques to understand complex, noisy, or non-linear data without being explicitly programmed (Tsangaratos and Ilia, 2016).

Machine learning algorithms are automated systems which can learn from the data and use that knowledge to quickly and accurately classify the unseen data. A number of studies show that machine learning algorithms can correctly classify a massive amount of data at a greater performance than human. For example, a study by Liu et al. (2017) uses Convolutional neural network (CNN) to distinguish cancer cells in medical slides. Convolutional neural network gives 92% of accuracy in cancer detection whereas human pathologist achieves only 73%. Moreover, the model spends seconds per checking one slide to diagnose cancer while human spends minutes. A another example is Wang et al. (2013). They use Artificial neural network (ANN) to classify well logging data from the Appalachian basin into the Marcellus shales which are defined by geochemistry analysis. By doing this, ANN gives 72% of average accuracy in well logging classification from the Appalachian basin. Consequently, machine learning is a powerful tool to improve data classification efficiently.

There are many machine learning algorithms and they have been developed for various data types such as text, picture, and video (Wu and Zhao, 2018). Each model is applicable to different data types. Hence, six machine learning algorithms: Support vector machine (SVM), K-nearest neighbour (K-nn), Extreme gradient boosting

(XGB), Artificial neural network (ANN), Convolutional neural network (CNN), and Gated recurrent unit (GRU) are employed to classify well logging data in this study in order to determine suitable algorithms for well logging classification. These models show high accuracy in lithological classification tasks in previous studies, e.g. Dubois et al. (2007); Hall and Hall (2017); Xie et al. (2018).

To compare the accuracy and efficiency of machine learning algorithms, well logging data of four geothermal wells, namely WO2, Mountain Home, Kimma, and Kimberly, from the Snake River Plain in Idaho, USA are used in this study. Each well contains approximately 4 - 12 logging types. The total number of logging tools or features is 15. The Snake River Plain is one of six high-grade geothermal areas in the US and it has high potential to become a geothermal energy source in the future (Tester et al., 2006). It is characterized by volcanic igneous rocks, in particular, basalt and rhyolite from extensional tectonic processes with being bound by normal fault (Bedrosian and Feucht, 2014).

In this study, the well logging data from four geothermal wells are split into three subsets: training data, validation data, and test data. Training data is used to train the models and validation data is used to optimize the model. After that, the models are evaluated by test data. There are two experimental designs in this study. The first scenario is designed to train (70% of data) and test (30% of data) the models with well-logs data from the same well. The second scenario uses a large amount of data as a training set, which combines three wells and 70% of the fourth well, and uses the remaining 30% of the fourth well as a test set. A comparison of model performance is further evaluated to determine the best and worst algorithm for lithological classification using well logging data from the Snake River Plain in two scenarios: single- and multiple-well tests. Furthermore, machine learning algorithms being automated system could improve efficiency and reduce time in well logging interpretation.

1.2 Objectives

- To classify lithology from well logging data by machine and deep learning algorithms
- To compare the result of each machine and deep learning algorithms

1.3 Scope of works

The objectives of this study to employ machine learning algorithms to classify well logging data from Snake River Plain as lithology and to compare the performance of each model. A total of six machine learning algorithms, namely SVM, K-nn, XGB, ANN, CNN, and GRU are used in this study due to give high performance in previous works (e.g. Dubois et al. (2007); Hall and Hall (2017); Xie et al. (2018)). The ANN, CNN, and GRU models are considered deep learning algorithms due to deeper structure compared to the others. Every model is trained and tested on the same data set with the same features scaling and features selection. The models are expected to classify well logging data into lithology such as basalt, rhyolite, and tuff based on well reports. Moreover, every rock type is equally important for this study. This is because the objectives of this project are to classify rock types generally and improve the efficiency of well logging interpretation. The objective does not include the determination of reservoir rock or oil bed. The models are measured by classification accuracy and confusion matrix in this study. Precision, recall, and F1 score is not used in this study.

Our data comes from two sources: Frontier Observatory for Research in Geothermal Energy and project HOTSPOT. The former project is led by Idaho National Laboratory (Podgorney, 2016a) whereas the latter is led by Utah State University (Shervais, 2014a, b, c). Frontier Observatory for Research in Geothermal Energy publishes WO2 well data drilled in 1991. Mountain Home, Kimma, and

Kimberly well are published by project HOTSPOT and drilled from 2010 to 2012. As four well logging data are gathered in different periods and by the different organizations, collected features, and lithological classification are disparate. For instance, resistivity log is not gathered in WO2 well. Thus, missing values or uncollected features must be tackled with some techniques when four well data are combined in scenario two. Furthermore, sedimentary rocks are classified into sandstone, siltstone, and claystone in WO2 well whereas they are grouped as a sedimentary rock in the other wells. Hence, rocks are reclassified in scenario two.

1.4 Expected results

- Lithological classification using well logging data from the Snake River Plain by deep and machine learning algorithms in single and multiple-well tests
- The optimal model for well logging classification using data from the Snake River Plain
- Classification performance of the optimal model with accuracy more than 70%

Chapter 2 Literature Reviews

2.1 Well logging

Generally, there are two techniques to gain information about subsurface rocks. The former is collecting subsurface rocks directly by drilling and the latter is interpretation indirectly from geophysical data. As drilling is expensive and time-consuming, the indirect method is more popular than the former. Well logging is a widely used geophysical method to gain information about subsurface lithology. This is because well logging can collect high-resolution data in millimeter to centimeter unit continuously (Liu, 2017). Consequently, well logging is used in this study in order to give information about the lithology of subsurface rocks. There are various types of logging tool which have been invented to collect different geophysical signal to infer lithologies such as gamma ray, resistivity, and neutron logs. In contrast, some logging tools do not measure the properties of rocks but it measures well properties, for example, caliper log (Fanchi, 2002). Hence, only logging tools referred to lithology and physical properties of rock are used in this study. Figure 1 shows an example of well logging data from the Anda Sag, the Songliao Basin (Huang et al., 2015).

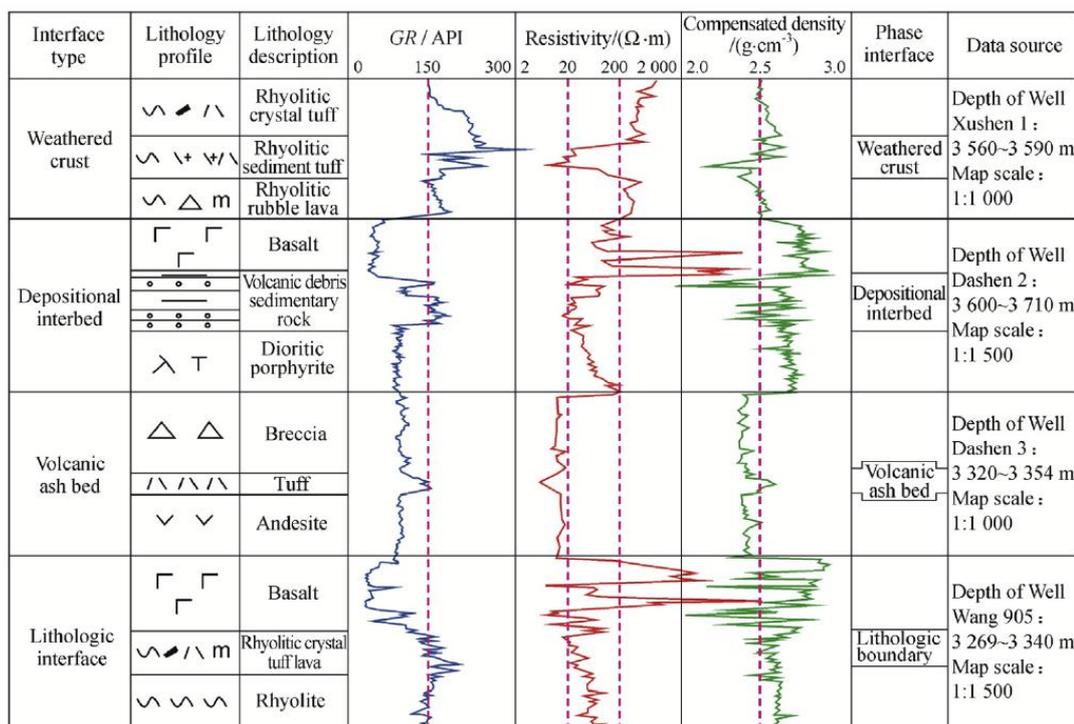


Figure 1 An example of well logging data of volcanic rocks in the Anda Sag, the Songliao Basin (Huang et al., 2015).

2.1.1 Resistivity log

Resistivity log applies the same principle of resistivity exploration. Resistivity is an inversion of conductivity so resistivity log measures the electrical resistance of subsurface rock. The resistivity of rock is mostly depended on the type of pore fluid. As water or brine contains a lot of ions, it has low resistivity. In contrast, the concentrations of ions in hydrocarbon is low. It has high resistivity (Fanchi, 2002). Moreover, there are three resistivity logs in this data set to give comprehensive results: deep measurement, shallow measurement, and mud measurement. The deep and shallow measurements represent resistivity of rock formation based on distance from logging tool while mud measurement represents resistivity of drilling mud. An example of a resistivity log is shown in Figure 1.

2.1.2 Gamma ray log

Gamma ray log is a passive logging tool which only receive a geophysical signal from subsurface rocks. Gamma ray log detects gamma ray emitted from radiometric elements in the rocks by Geiger-Mueller (G-M) counter or the Scintillation counter (Liu, 2017). The API is standard unit for gamma ray, where one API represents about 0.07 micrograms of radium equivalent per ton of formation (Tiab and Donaldson, 2012). In general, potassium, thorium, and uranium are major sources of the natural gamma ray in subsurface rock (Tiab and Donaldson, 2012). As clay minerals are composed of these three elements, shale emits more gamma ray than other rocks such as sandstone, limestone, and dolostone. Gamma ray is normally used to distinguish between shale and other rocks and used to calculate the fraction of shale in a rock formation. Moreover, gamma ray log can detect sources of a gamma ray from in rock formation by gamma ray spectrum in order to quantify the volume of potassium, thorium, and uranium in the rock formation (Liu, 2017). Figure 2 provides the correlation between the gamma ray log and sedimentary section (Martinius et al., 2002).

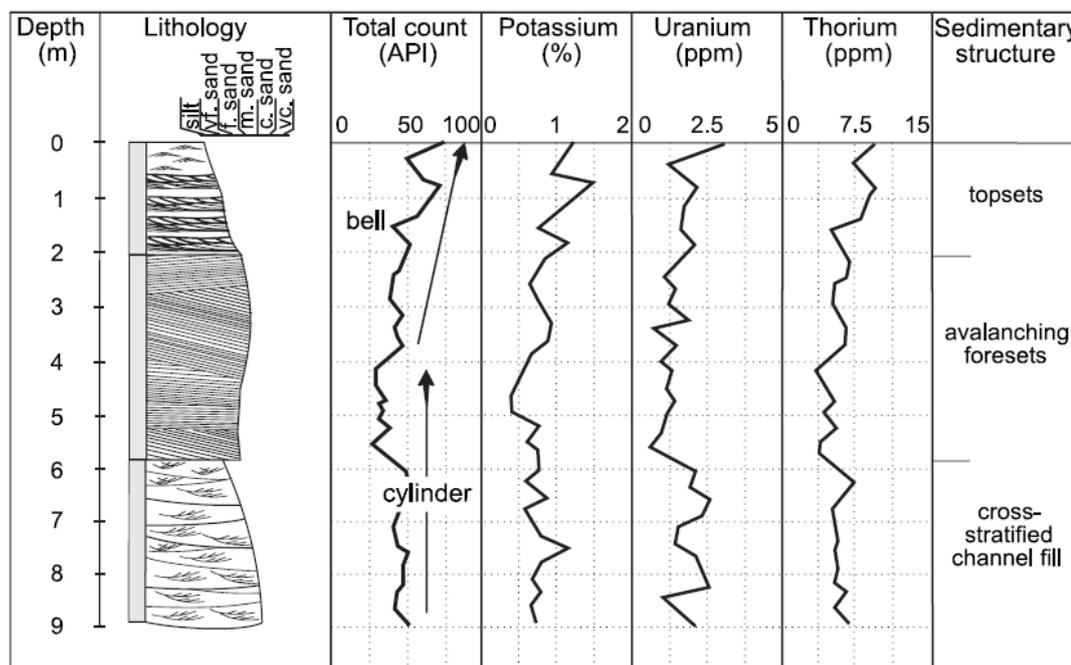


Figure 2 An example of gamma ray log correlated with the sedimentary rock section in Loranca Basin, Spain (Martinius et al., 2002).

2.1.3 Acoustic log

Acoustic log or sonic log is active logging tools which send energy and receive the geophysical signal from the subsurface rock. This logging tool sends a sound wave to the subsurface rock and records travel time of the sound wave passing through the rock (Fanchi, 2002). The sound wave is a function of density and elastic properties of the medium and different rocks have different density and elastic properties. Moreover, porosity and pore fluid affect density too. Consequently, the acoustic log is generally used to determine fluid content, porosity, and lithology (Liu, 2017). Figure 3 presents a comparison between gamma ray log and acoustic log (V_p and V_s) of the upper and lower Marcellus shale and Cherry Valley limestone (Roy, 2013).

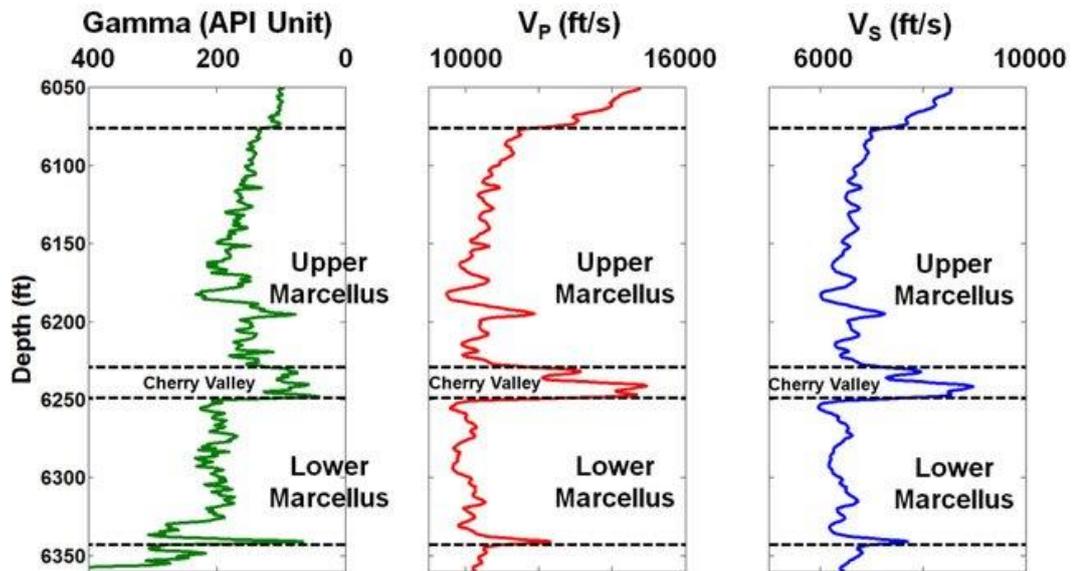


Figure 3 The well logging data, gamma ray log (solid green line), P-wave sonic log (solid red line), and S-wave sonic log (solid blue line) of the upper and lower Marcellus shale and Cherry Valley limestone (Roy, 2013).

2.1.4 Density log

Density log or photoelectric log measures density of subsurface rock by Compton scattering. Gamma ray from density log is injected to subsurface rock and it loses some energy to an atom of the rock formation. After that density log measures the remained gamma ray from the rock formation. The intensity of gamma ray arriving in density log is a proportion of rock density (Tiab and Donaldson, 2012). Normally, density log is useful to determine porosity. Moreover, the type of fluid affects density too since density log can infer to pore fluid in the rock formation. An example of a density log is shown in Figure 4.

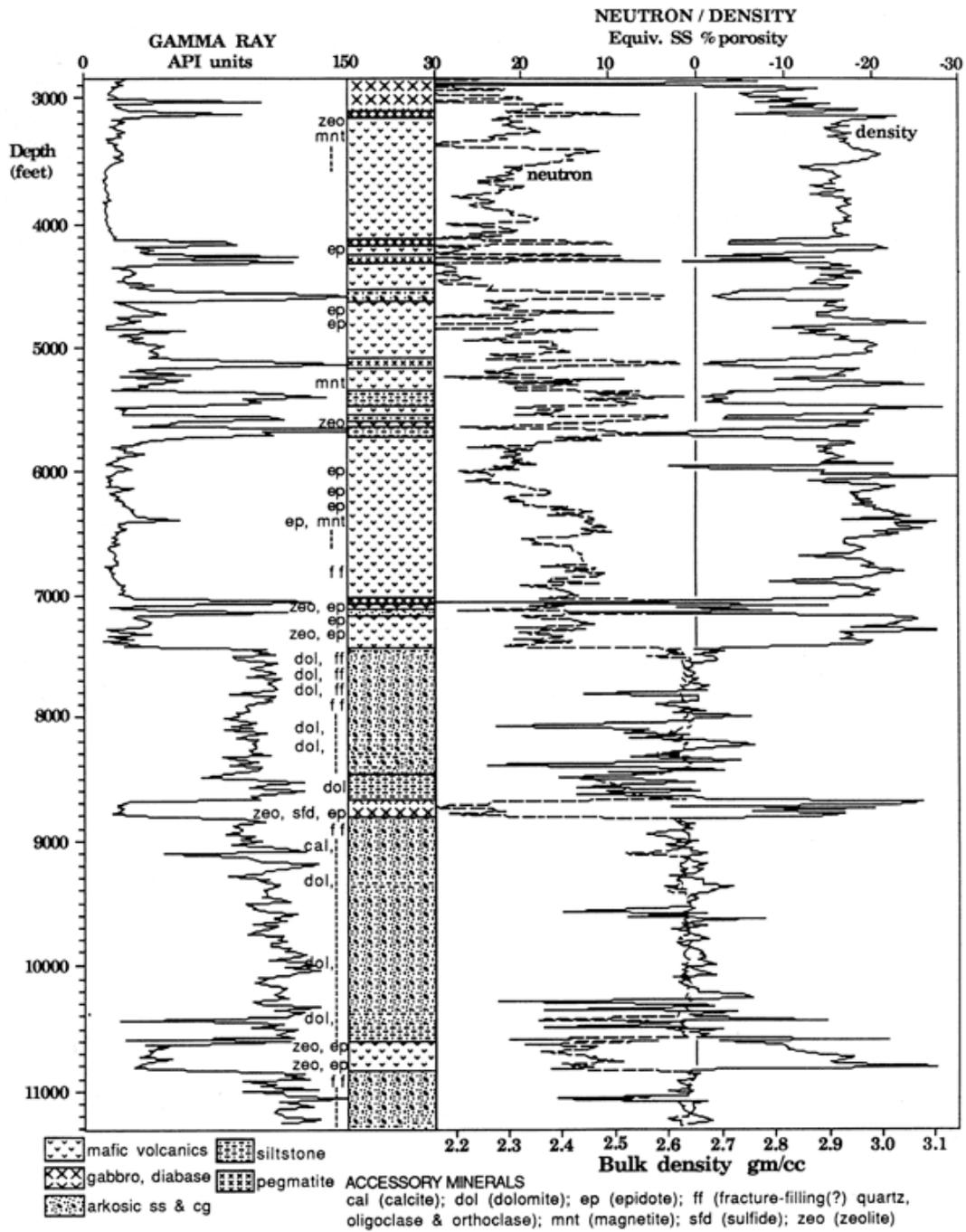


Figure 4 A correlation of well logging data, gamma ray log, density log, and neutron log, with lithology in the Rift System, Kansas (Berendsen et al., 1988).

2.1.5 Neutron log

Neutron log emits continuous fast neutron to subsurface rock. The fast neutron is slowed down while releasing secondary gamma ray by collision with a hydrogen atom from the subsurface rock. Hydrogen content of rock formation can be evaluated by counting slow neutron or secondary gamma ray (Liu, 2017). The assumption is that the major source of hydrogen in subsurface rock is liquid in pores. Thus, hydrogen content refers to porosity. Moreover, neutron log presents very low density when pores are filled with gas. This is because gas has a low number of atoms. Thus, neutron log can be used to detect the type of pore fluid (Tiab and Donaldson, 2012). However, rocks which have high hydrogen content such as coal and organic shale have high influence with neutron log. Consequently, the neutron log has to work with other logs to reduce the uncertainty of measurement. Figure 4 presents well logging data of gamma ray, density, and neutron logs of the Rift System, Kansas (Berendsen et al., 1988).

2.1.6 Direct measurement log

Some logging tools measure information about borehole environment such as driller's Logs, caliper Logs, sample Logs (Fanchi, 2002). In this study, temperature and pressure logs are used in well logging classification because our data comes from geothermal wells. Temperature log measures formation temperature in Celsius degree and pressure log measures formation pressure in Pascal (Fanchi, 2002).

2.2 Machine learning algorithms

2.2.1 Previous works

Machine and deep learning algorithms are applied in various fields, for example, medical examination, morphology, and climate modeling. A study by Catalogna et al. (2012) applies ANN to monitor glucose in rats. Results suggest that ANN can predict glucose level accurately, followed by errors being below than 10%. In addition, a study by López et al. (2017) compares the efficiency between ANN and mathematic equation from previous works in point bars modeling in Spain. This work aims to predict a position of point bars from previous data such as the slope of the beach, the steepness, and satellite data. ANN exhibits higher accuracy score about 10 - 20% than the proposed methods. This is because ANN has more generalization than the proposed methods. ANN can learn the unique characteristics of point bars in Spain from the previous data while the previous equations for calculating point bar position are generated from other areas. Consequently, the position of point bars from ANN is closer to the reality that the position of point bars from the proposed methods. Another study by Devak et al. (2015) uses SVM, K-nn and a hybrid model between SVM and K-nn to predict precipitation in India. This study shows that the hybrid model provides higher performance than the SVM and K-nn.

In well logging, there are a number of studies applying machine learning to improve the accuracy of lithological classification. A study by Wang et al. (2013) uses ANN to classify Marcellus shale into organic siliceous shale, organic mixed shale, organic mudstone, gray siliceous shale, gray mixed shale, gray mudstone, and carbonate intervals. The shales are defined by core analysis. Results show that ANN gives average accuracy in well logging classification at about 73%. Although 73% is mediocre results, shale classification using well logging data surpasses the ability of human. This is because, in general, geologist classifies well logging data into sandstone, shale, or limestone. The shale facies in Wang et al. (2013) come from

geochemistry results. Xie et al. (2018) applies four machine learning algorithms, Naïve Bayes, SVM, ANN, and XGB, to classify well logging data from Ordos basin in China. Results show that XGB provides the highest accuracy at about 83% while Naïve Bayes show lowest accuracy (<50%). As Naïve Bayes requires assumption that features are conditionally independent, it is unsuitable with well logging data which features are dependent.

As different machine learning algorithms give different results, various models are employed in this study in order to determine the best algorithm for well logging data from Snake River Plain. XGB is selected because it performs well in a study by Xie et al. (2018). SVM and K-nn give comparable efficiency in Devak et al. (2015) and SVM gives acceptable accuracy score in Xie et al. (2018). Hence, they are employed in this study. ANN is one of deep learning algorithms. Even though ANN shows the low accuracy in Xie et al. (2018), it exhibits high accuracy in other tasks such as glucose monitoring (Catalogna et al., 2012), point bars prediction (López et al., 2017), and shale classification (Wang et al., 2013). Consequently, ANN is applied in this study. Moreover, CNN and GRU which are other deep learning algorithms are employed in this study to evaluate whether deep learning algorithms are suitable with well logging data.

2.2.2 Support vector machine (SVM)

The idea of Support vector machine (SVM) classification is to define the decision boundary between each class from training dataset and to use this boundary to classify data in the test dataset. However, there are various decision boundaries for one data set. SVM choose the best decision boundary based on margin. It is the twice distance between decision boundary and the closest data point. The optimal decision boundary for SVM is the line having a maximum margin so the goal of SVM is to maximize margin (Xie et al., 2018). The C parameter is tuning

parameter for soft margin which accepts some error to generalize the model (Cristianini and Taylor, 2000). It indicates how much SVM cares about the error (Figure 5). If C is large meaning that SVM cares so much about the error, SVM will try to reduce the number of misclassified data points. In this study, radial basis function kernel (RBF) is used to solve non-linear task for SVM (Xue et al., 2009). The concept of RBF is putting Gaussian onto every data point (Figure 6), while γ manage how decision surface spread. If γ is high, the decision surface will be dramatically curved. As SVM with soft margin and RBF kernel is used in this study, there are two parameters modified, namely C and γ .

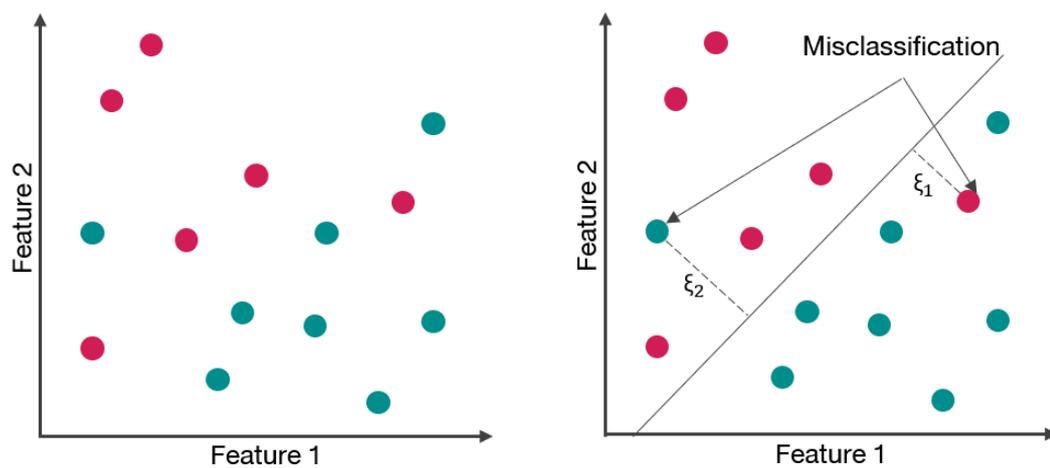


Figure 5 A case in point of SVM with soft margin. Right represents data set which has noisy data. Left shows separating hyperplane divided classes with having some misclassified data points.

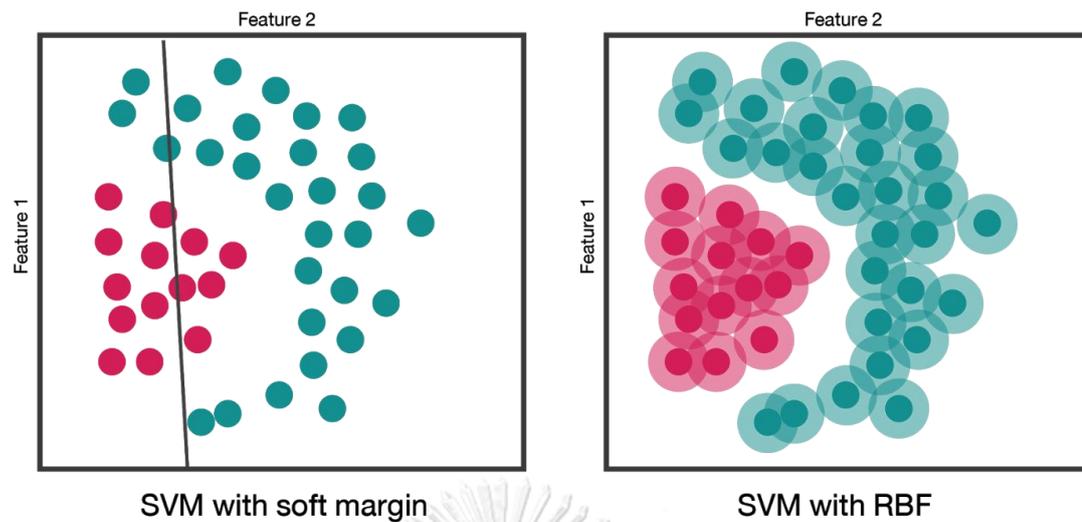


Figure 6 The concept of SVM with RBF. Left is SVM with soft margin and black line is decision boundary between Class 1 and Class 2 . Right is SVM with RBF. Pink and green surface are decision boundary between Class 1 and Class 2.

2.2.3 K nearest neighbour (K-nn)

K nearest neighbour (K-nn) is a simple deep learning algorithm used for widely both classification and regression tasks (Glowacz and Glowacz, 2016). This model finds K nearest training data points around the test data point and then the class of the test data point is defined by the most popular class among K nearest data points (Rastegarzadeh and Nemati, 2018). K-nn is lazy learning which delays generalizing training data until the query arrived, as opposed to eager learning which generalizes training data before the query received (Steinbach and Tan, 2009).

First, training data set which is labeled and testing data point which is unlabeled are given to K-nn (Figure 7). As a next step, the distance between each of training data points and test data point is calculated to find K nearest data points and in this case, K is equal to 3 that means K-nn will find 3 closest data points around testing data point. Then unlabeled data point is assigned class based on the most popular class.

Generally, there are three elements to modify this approach: the number of nearest neighbours (K), distance function, and the voting method (Steinbach and Tan, 2009). However, the value of K affects the most to a performance of K -nn for a complex task. Too large or too small K value lead the model to the wrong classification as shown in Figure 8. In contrast, the optimal neighbourhood is led it to reduce the complexity of the model and increasing its accuracy (Shen et al., 2016).

K-Nearest Neighbor

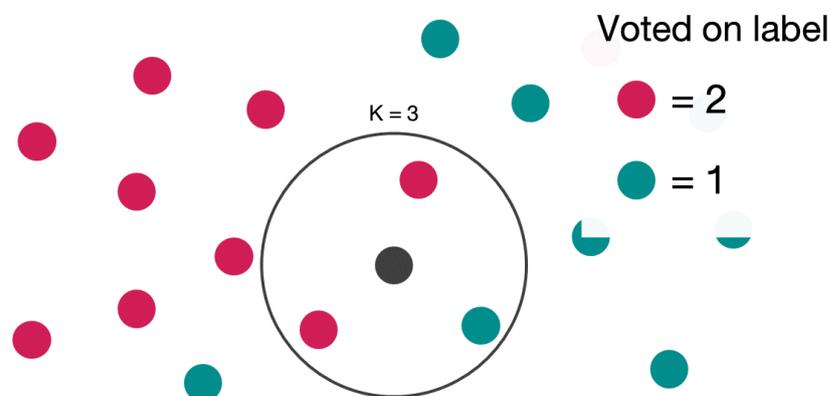


Figure 7 Classification of unlabeled data point by majority vote. The unlabeled is classified to Class 1 because Class 1 is the most popular class among the K nearest data points.

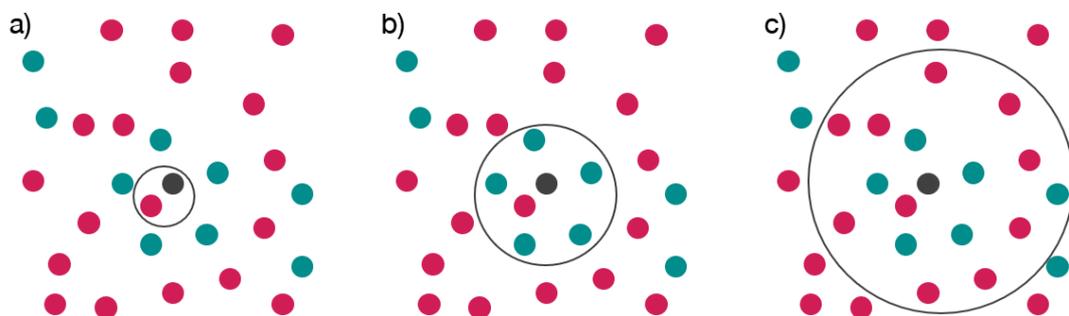


Figure 8 K nearest neighbour with too small K (a), suitable K (b), and too large K (c).

2.2.4 Extreme gradient boosting (XGB)

Decision tree is a foundation of Extreme gradient boosting (XGB). It iteratively partitions input data from training dataset into subdivisions based on their features until every class of training dataset is clearly separated from each other (Friedl and Brodley, 1997). For example, training data comprise of sandstone (pink) and shale (green) and there are two features: volume of clay and grain size. First, in the root node, data are split into three groups by volume of clay and then for some group, which two classes are still mixed, decision tree uses the other feature to separate shale from sandstone. Finally, this condition (decision tree) is applied to classify unlabeled data point (black) as shown in Figure 9.

However, a single decision tree is not strong enough to handle sophisticated data, in particular, on the data having several features and classes. As a result, ensemble method which combines several decision trees together is proposed in order to ameliorate accuracy and reduce bias. This approach is divided into averaging method and boosting method (Xie et al., 2018). The concept of averaging method is independently to construct decision trees and average their voted results. A case in point is Random Forest introduced by Breiman (2001). By comparison, boosting method aims to construct decision trees sequentially to reduce bias of model. XGB is one of boosting ensemble method.

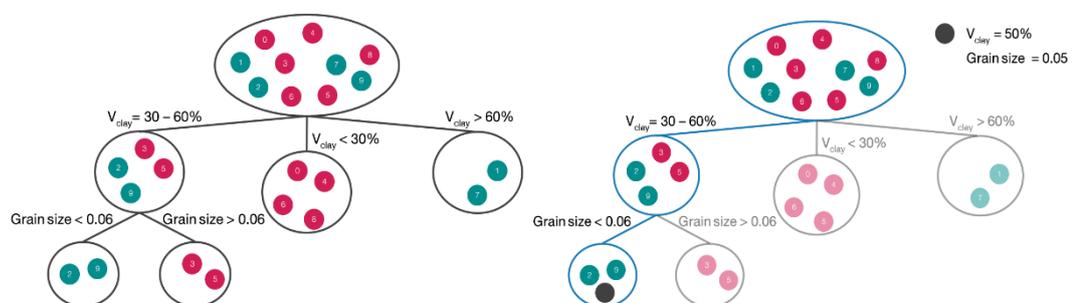


Figure 9 The procedures of decision tree. First, training data are used to construct a decision tree and then it is applied to classify new data point from test data.

Learning rate or shrinkage is the parameter which determines how much the model is adjusted or changed based on the loss function (Friedman, 2002). Too small learning rate lead model to overfitting and getting stuck in local minima. In contrast, too large learning rate encourages model to underfit. Min child weight is the minimum sum of instances weigh in leaf. This parameter uses to avoid overfitting too. If sum of instances weigh in leaf is lower than min child weight, the leaf will be splitted. In simple word, it inhibits model to over-reliance on one potential feature to prevent overfitting. Max depth defines the maximum number of splitting in each tree. This parameter refers to complexity of model. High max depth make model more complex and specific to particular one data set so it causes our model overfitting. However, too low max depth leads model to underfitting.

Every parameter controls generalization of our model because XGB has ability to fit every data point in training data set to obtain 100% accuracy in training data set. However, this does not mean model will fit in every data point in testing set and it usually acquires low score of accuracy in testing data set. In contrast, model can find simple relationship like linear or curve line between data and label but it is not suitable relationship which explains our data set. The concept of overfitting and underfitting are shown in Figure 10. Therefore, changing parameter is required to increase predictive accuracy while generalize our model. More details about XGB including mathematical equation and tuning parameters can be found in Chen and Guestrin (2016).

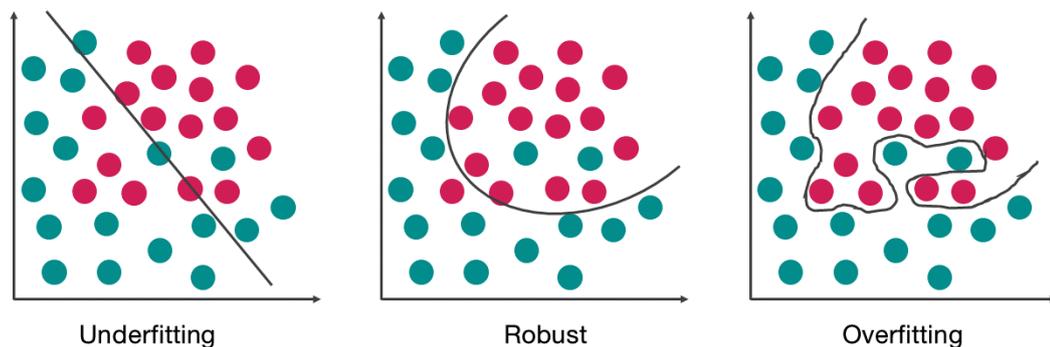


Figure 10 The concept of overfitting and underfitting. On the left, model cause underfitting so decision boundary

2.2.5 Artificial neural network (ANN)

A study by Mcculloch and Pitts (1943) first introduces the simplest neural network, which was inspired by the complex network of human brain. In that time, the efficiency of computer was not high. Moreover, other models which do not require high performance computer and storage like K-nn, SVM, and decision tree was announced so neural network was not popular at that time. Then at the end of 20th century, computer have been develop until it can deal with tremendous data with complex models (Yadav et al., 2015). As a result, neural network begins more famous and it is renamed to Artificial neural network (ANN).

ANN is deep learning algorithm inspired by information processing of biological system comprised neurons and complex connection system between these neurons (Bishop, 2006). ANN consists of three layers: input layer, hidden layer, and output layer. As a analogy between ANN and neuron (Figure 11), input layer is similar to dendrite which receives data and then passes it to cell body. Hidden layer resembles the cell body which is responsible for determining a relationship between input data and generating the label of each data. Moreover, ANN can have more than one hidden layer to improve the performance of the model. This study uses three hidden layers for ANN and other neural network models. The last one is output

layer which is equivalent to axon, which aims to transmit the information from cell body to another neuron or muscle (López et al., 2017).

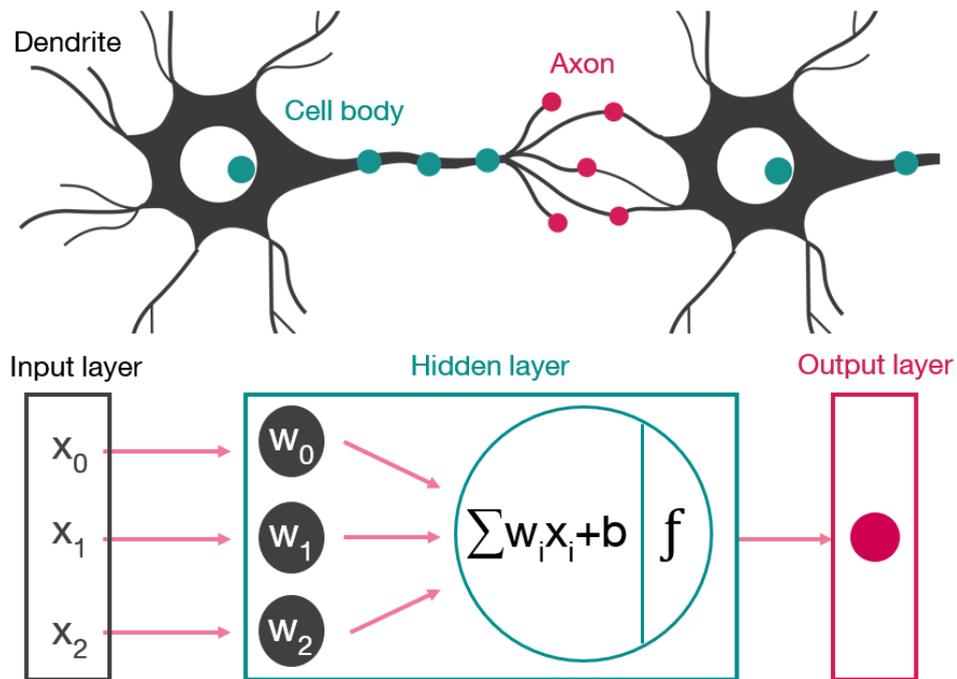


Figure 11 A analogy between nerve system (top) and Artificial neural network (bottom).

Sometime when neural network which has a lot of nodes works on small dataset, the results seem likely to be poor. As the number of nodes and data are unsuitable, neural network is inclined to too use some specific nodes. This would lead neural network to overfitting with training data. To prevent this overfitting, dropout is applied to neural network. It omits some nodes in each iteration of training step to prevent the model to over-reliance on some specific nodes (Srivastava et al., 2014). It is set in term of probability. For example, 0.5 means that half of nodes are dropped every iteration in training phase (Figure 12). By doing this, neural network will not over-reliance on specific nodes and increased generalization of model.

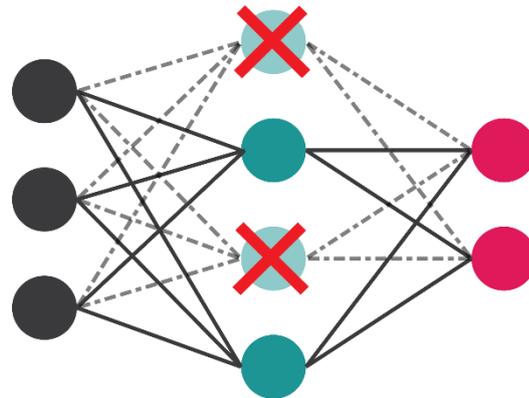


Figure 12 Dropout parameter. Dropout eliminates some of nodes while training step to prevent model to over-reliance on some nodes.

2.2.6 Convolutional neural network (CNN)

Convolutional neural network (CNN) is subset of Artificial neural network (ANN) which adopts convolution layers. It was proposed by Cun et al. (1989) to apply in data containing many features. Then the various types of CNN have been created like VGG (Simonyan and Zisserman, 2014), GoogLeNet (Szegedy et al., 2015), and ResNet (He et al., 2016) with providing a high efficiency in image classification and detection tasks (Wu and Zhao, 2018). In ANN, there are 3 main layers consisted of input layer, hidden layer, and output layer for classification. In contrast, CNN adds convolution layer and pooling layer which aim to extract certain features to turn raw data into representation which noises and unnecessary features are removed (Figure 13). After that, the representation is passed through to hidden and output layers respectively for classification step. By doing this, the number of features processed by network has been reduced and made network more efficient to implement.

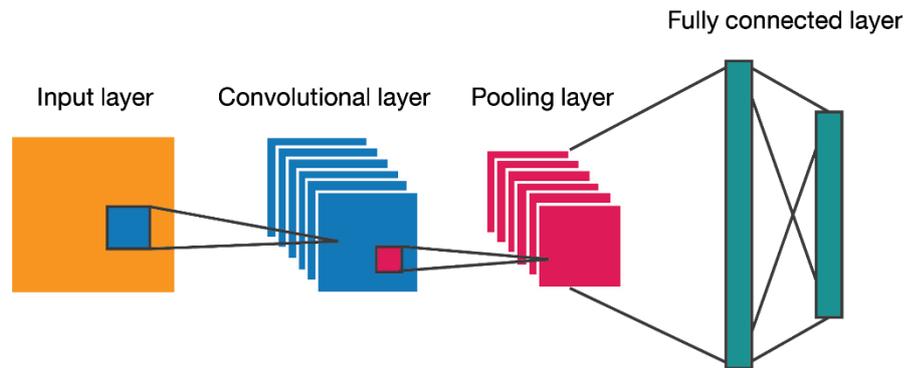


Figure 13 The structure of CNN. It consists of input layer, convolutional layer, pooling layer, and fully connected layer.

Convolutional layer is first layer of CNN which aims to extract only important features from input by convolutional operation. Convolutional operation refers to mathematical operation which combines two metrics by dot product and sums the result to produce new metric as shown in Figure 14 (Wu and Zhao, 2018). Convolution layer works as sliding each filter over the input and produces feature map. The network then correlates relationship between feature map and output. Filter is changed iteratively until the optimal filter which provide feature map having strongest relationship with output has been found. The proper filters will be learned by training procedure and are applied on test data for feature extraction.

Pooling layer is to reduce space of output feature map by stepping a stride over the feature map as convolution layer. There are two general functions for pooling layer: average and max operation but this study uses max operation which computes the max value of each feature map stepped by a stride. After data are passed convolutional and pooling layer, it will be sent to fully connected layer or ANN. Then optimal decision boundary is calculated by ANN.

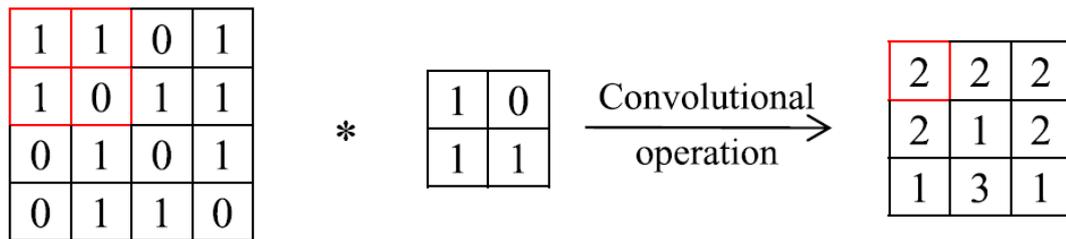


Figure 14 Convolutional operation.

2.2.7 Gated recurrent unit (GRU)

Gated recurrent unit (GRU) is a one of Recurrent neural network (RNN) which is suitable for sequential data (Figure 15). In CNN, our network remembers data in space series to extract important feature before pass through fully connected layer. In comparison, RNN is constructed to memorize data in time-series. It uses few previous and next time step data with the current data point as an input data to predict output. By doing this, RNN have been perfectly suitable to problem involved sequential data in particular speed recognition (Liu et al., 2018). To train RNN, the backpropagation through time (BPTT) is applied but it lead to cause the gradient vanishing problem when the network remember a long time series data (Bengio et al., 1994). Therefore, GRU which have reset and update gate to prevent the gradient vanishing is proposed (Liu et al., 2018). The former gate is designed to manage how much stored information which should be remove. The latter aims to consider what should be memorize and how it is important (Tutubalina et al., 2018).

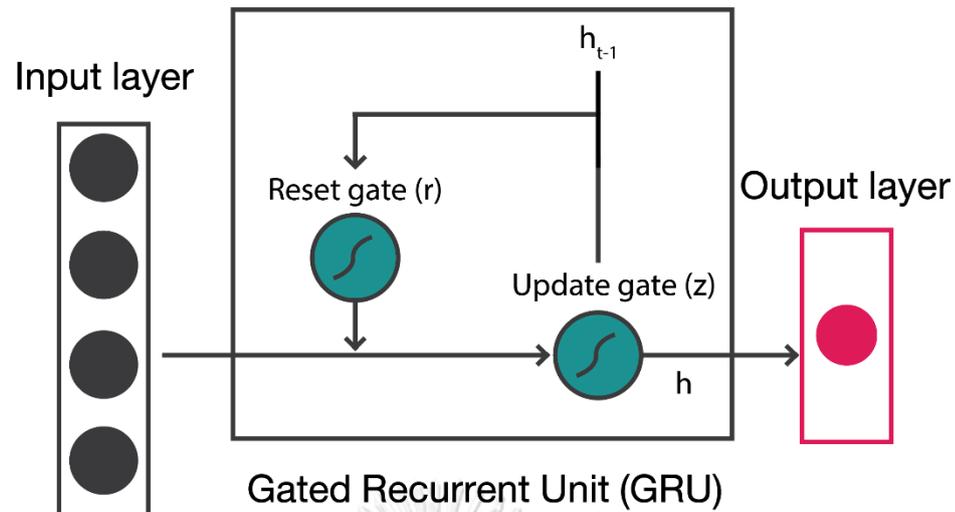


Figure 15 The conceptual model for GRU.

2.3 Experimental design

Experiments can be divided into two major groups: single well and multiple wells test. The former means that well logging data from one well are split into training and test data. For example, Xie et al. (2018) applies four machine learning algorithms on well logging data from the Ordos basin in China. Xie et al. (2018) has two datasets from two fields: Daniudui (five wells) and Hangjinqi (seven wells) gas fields. Machine learning algorithms are used to classify each dataset separately. To illustrate, data from five wells in Daniudui gas field are randomly split into training data (80%) and test data (20%). Then models are trained and tested with data from Daniudui gas field. As the next step, these procedures are repeated with data in Hangjinqi gas field.

In comparison, the latter means that many wells are combined as training dataset. Hall and Hall (2017) uses data from nine wells from Hugoton and Panoma fields, USA, in well logging classification task. One well is assigned as test dataset whereas eight wells are assigned as training dataset. After that, the models are trained on training data and tested on test data. In this study, the models are employed under two experiments: single well and multiple wells.

Chapter 3 Study Area

3.1 Overview of data

Well logging data and well reports from four geothermal wells (WO2, Mountain Home, Kimma, and Kimberly) in the Snake River Plain are used in this study. The location of the four wells is shown in Figure 16. The depth in each well varies from 2000 m to 4000 m. Each well has over 8,000 data points. This four wells data comes from two projects. The former is the Frontier Observatory for Research in Geothermal Energy (FORGE) which is led by Idaho National Laboratory (Podgorney, 2016b). The latter is project HOTSPOT which is led by John W. Shervais, Utah State University (Shervais, 2014a, b, c).

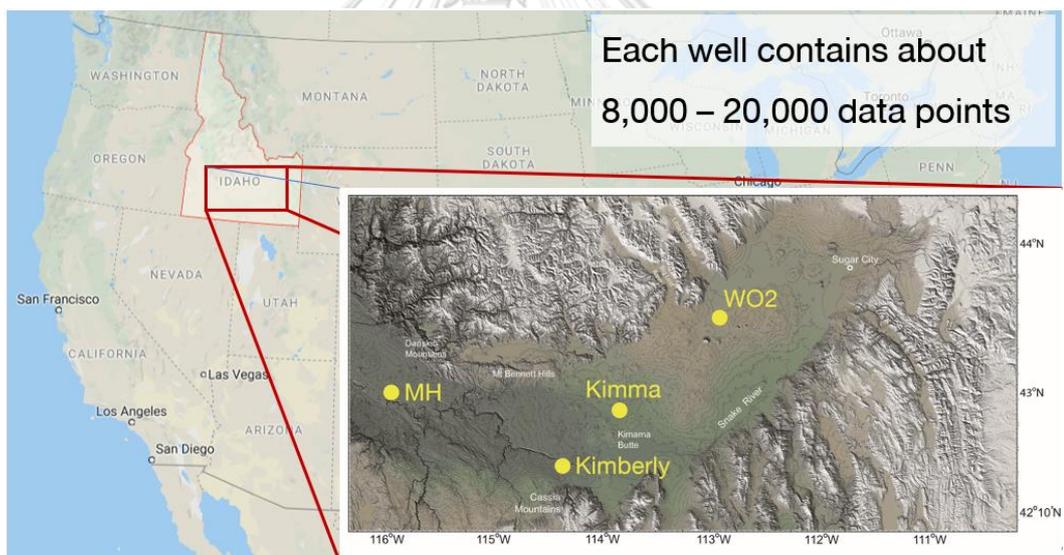


Figure 16 Location of 4 study wells. Yellow spots represent the location of study well. WO2, Kimma, and Kimberly well stand on Eastern Snake River Plain while Mountain Home locate on Western Snake River Plain (Shervais et al., 2013).

FORGE was established in 2015. They publish WO2 well data. FORGE aims to analyze geological data to enhance conceptual geological model in Snake River Plain. Consequently, they publish every data, for example, well logging data, core analysis, and hot water analysis. WO2 well was drilled in 1991 for about 5 km (Anders et al., 2014; Mazurek, 2004). In this well, there are four features: gamma ray, neutron, porosity, and density. Rocks in this well are classified into basalt, claystone, sandstone, siltstone, conglomerate, tuff, and vitrophyre.

The data from Mountain Home, Kimma, and Kimberly are collected in order to evaluate geothermal potential in three disparate geological sites (Shervais et al., 2013). These three wells are drilled between September 2010 and January 2012 while collecting for further researches (Shervais et al., 2013). Mountain Home is in a fault-bounded basin with thick sedimentary rock interbedded. Kimma is high sub-aquifer with a mafic intrusion. Kimberly is valley-margin with up-flow hot fluid. However, the data from some part of the whole well is used in this study. The rocks are classified into basalt, sedimentary rock, and rhyolite for these three wells and there are twelve features: gamma ray, temperature, pressure, resistivity in mud measurement (R_{mud}), resistivity in deep measurement (R_d), resistivity in shallow measurement (R_s), p-wave (V_p), s-wave (V_s), and water wave speed (V_w).

A total of fifteen types (features) of well logs and nine rock types (classes) are used to collect data. The types and measurement resolution of logging tools as well as lithological classification are shown in Table 1 and 2, respectively. The resolution of each logging tool is between 0.1 to 0.5 m. The distribution of classes in each well is presented in Figure 17. The vast majority of data is igneous rocks such as basalt, rhyolite, and tuff whereas sedimentary rocks are the minority. Moreover, Figures 18 to 21 provide the information about well logging data with lithology. An example of well logging data used in this study is shown in APPENDIX K.

Table 1 Available well-logs (features) in WO2, Mountain Home (MH), Kimma, Kimberly in the Snake River Plain.

Type of well-log (Features)	Well name				Resolution range (cm)
	WO2	Mountain Home	Kimma	Kimberly	
Gamma ray	✓	✓	✓	✓	20
Temperature		✓	✓	✓	1 - 5
Pressure		✓	✓	✓	1 - 5
Rmud		✓		✓	80
Rd		✓	✓	✓	80
Rs		✓	✓	✓	80
Thorium		✓		✓	20
Uranium		✓		✓	20
Potassium		✓		✓	20
Vp		✓	✓	✓	50
Vs		✓	✓	✓	50
Vw		✓			50
Density	✓				50
Porosity	✓				50
Neutron	✓				50

Remarks: Rmud refers to Resistivity in mud measurement. Rd refers to Resistivity in deep measurement. Rs refers to Resistivity in shallow measurement. Vp refers to p-wave. Vs refers to s-wave and Vw refers to water wave speed.

Table 2 Rock types (classes) in WO2, Mountain Home (MH), Kimma, Kimberly in the Snake River Plain.

Rock type (Classes)	Well name			
	WO2	Mountain Home	Kimma	Kimberly
Basalt (BS)	✓	✓	✓	✓
Claystone (CS)	✓			
Conglomerate (CG)	✓			
Sandstone (SS)	✓			
Siltstone (ST)	✓			
Tuff (TF)	✓			
Vitrophyre (VP)	✓			
Sedimentary rock (SR)		✓	✓	✓
Rhyolite (RH)				✓

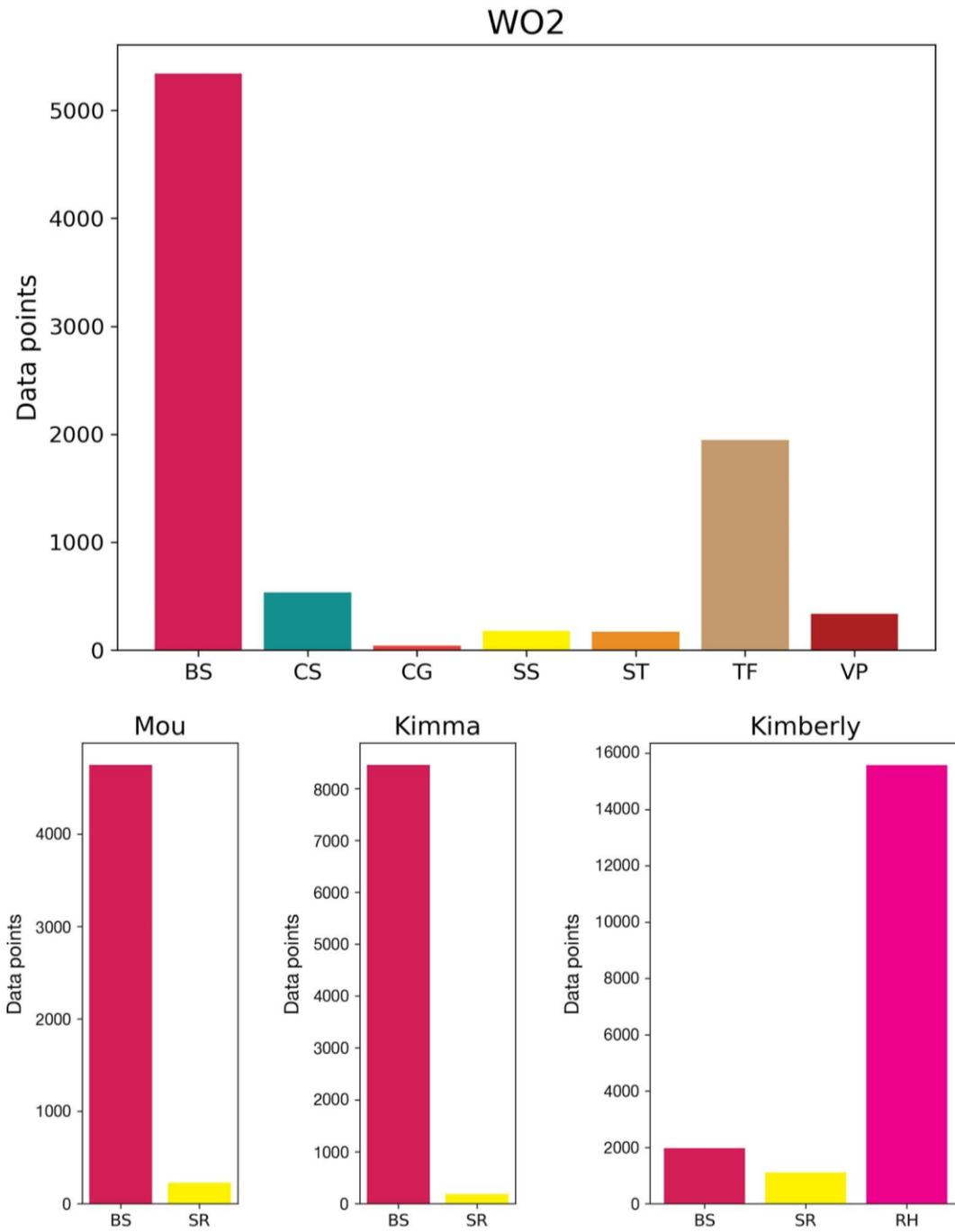


Figure 17 Classes distribution of each well. BS: basalt, CS: claystone, CG: conglomerate, SS: sandstone, ST: siltstone, TF: tuff, VP: vitrophyre, SR: sedimentary rock, and RH: rhyolite.

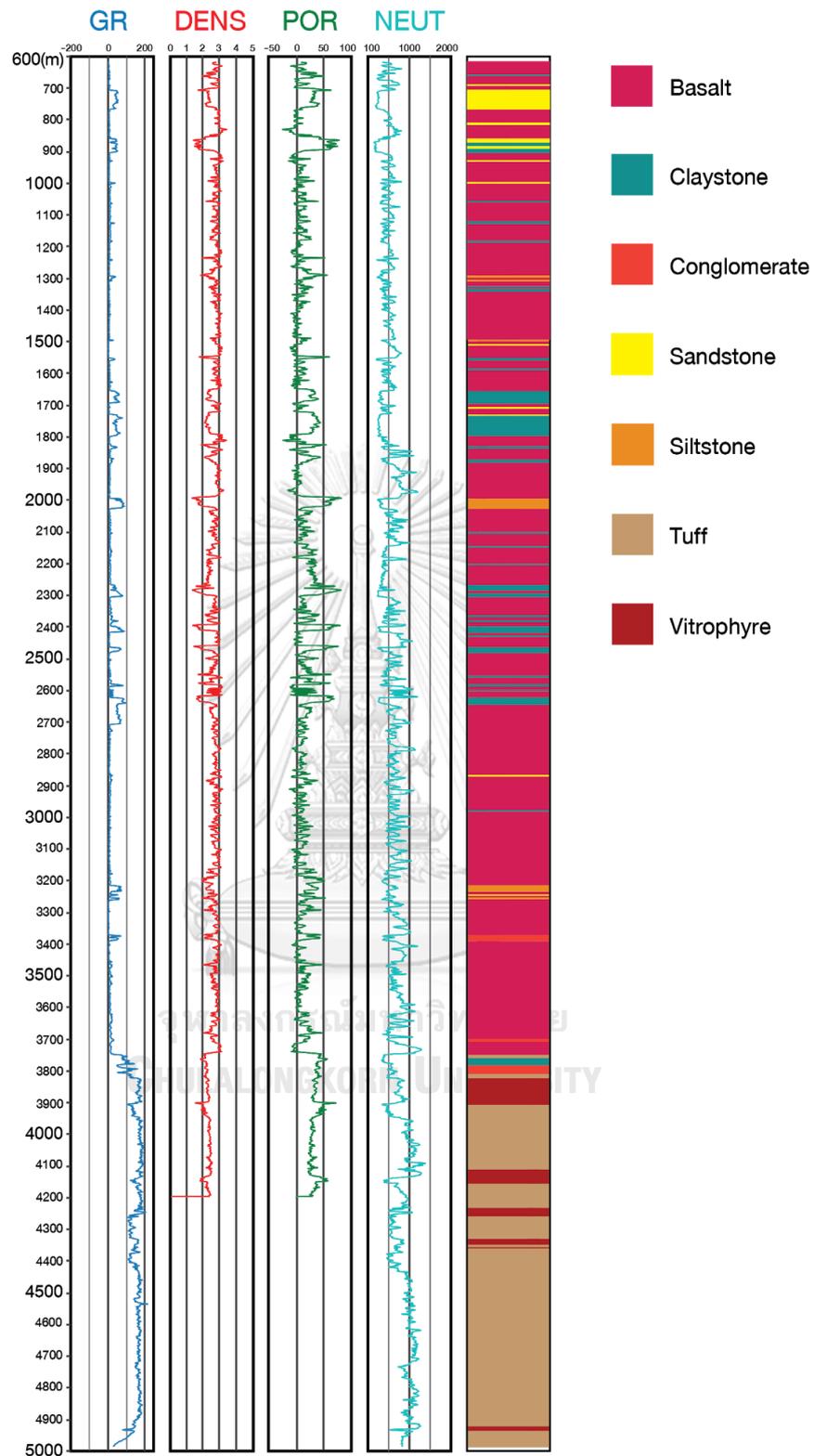


Figure 18 Well logging data of WO2 with lithostratigraphic log. This well depth is about 5000 m with 0.5 m interval. GR: gamma ray, DENS: density, POR: porosity, and NEUT: neutron.

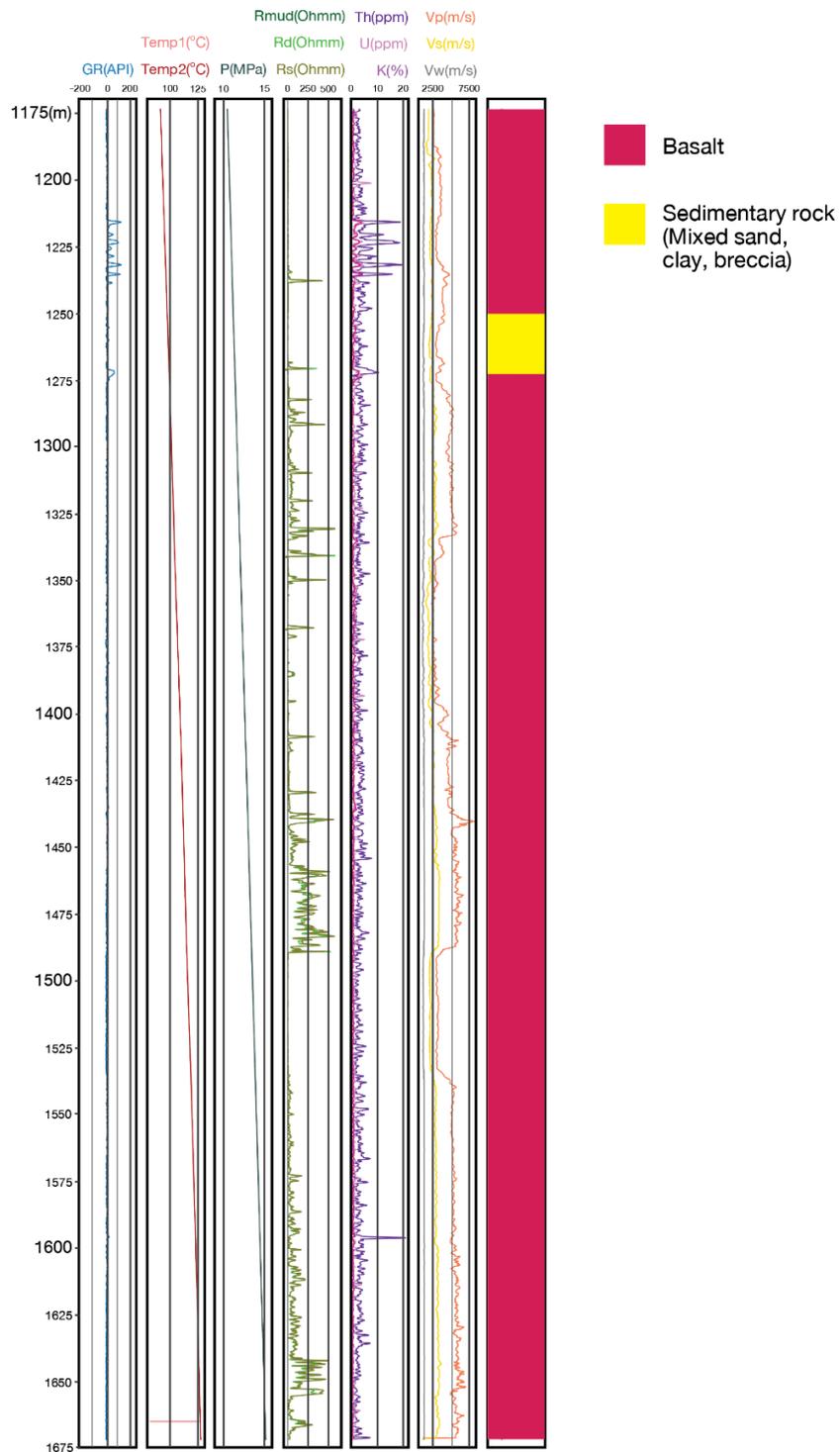


Figure 19 Well logging data of Mountain Home with lithostratigraphic log. GR: gamma ray, Temp: temperature, P: pressure, Rmud: resistivity in mud measurement, Rd: resistivity in deep measurement, Rs: resistivity in shallow measurement, Th: thorium, U: Uranium, K: Potassium, Vp: P-wave velocity, Vs: S-wave velocity, Vw: water wave velocity.

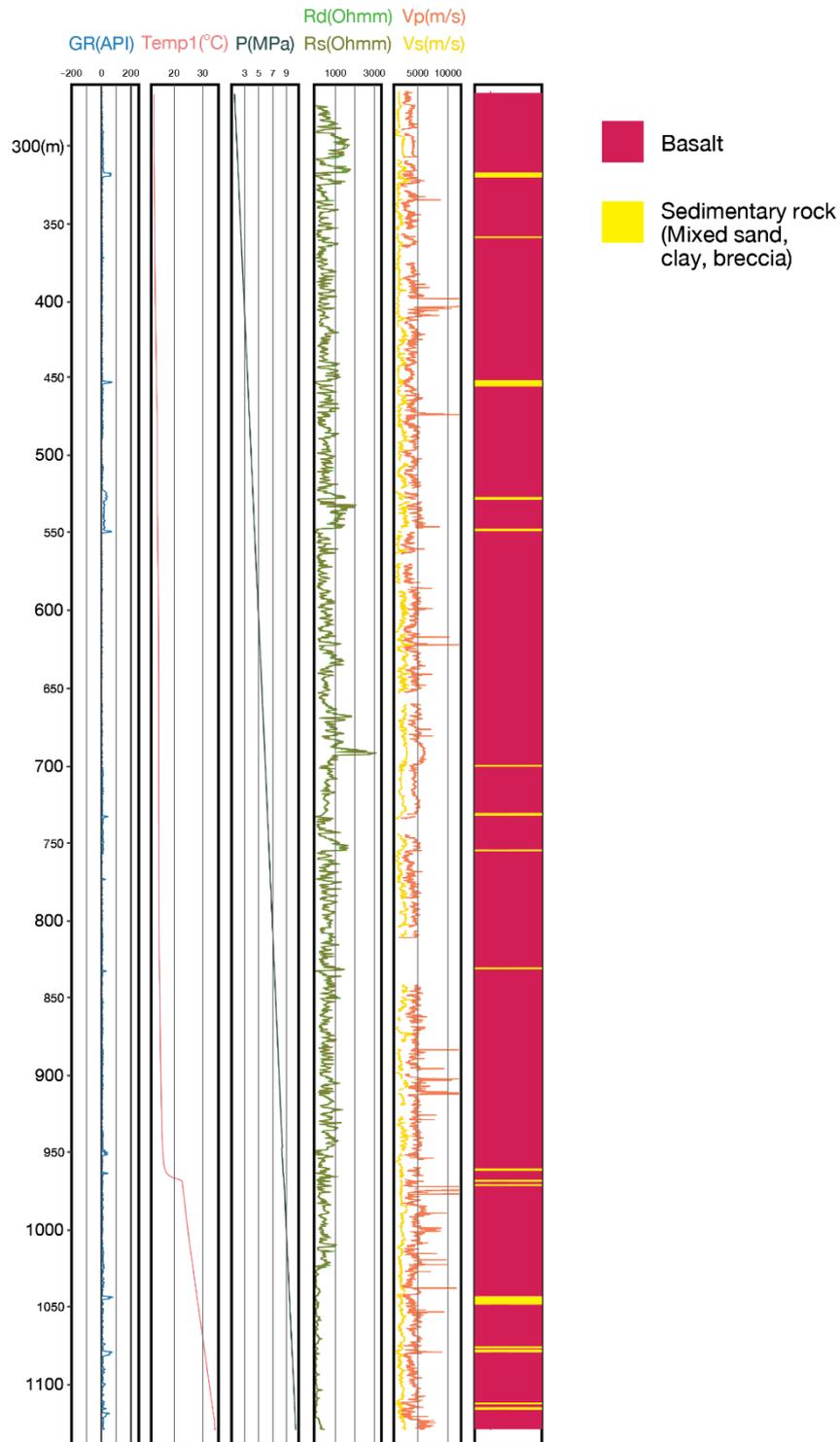


Figure 20 Well logging data of Kimma with lithostratigraphic log. There are about 800 m in Kimma well depth with 0.1 m interval. GR: gamma ray, Temp: temperature, P: pressure, Rd: resistivity in deep measurement, Rs: resistivity in shallow measurement, Vp: P-wave velocity, Vs: S-wave velocity.

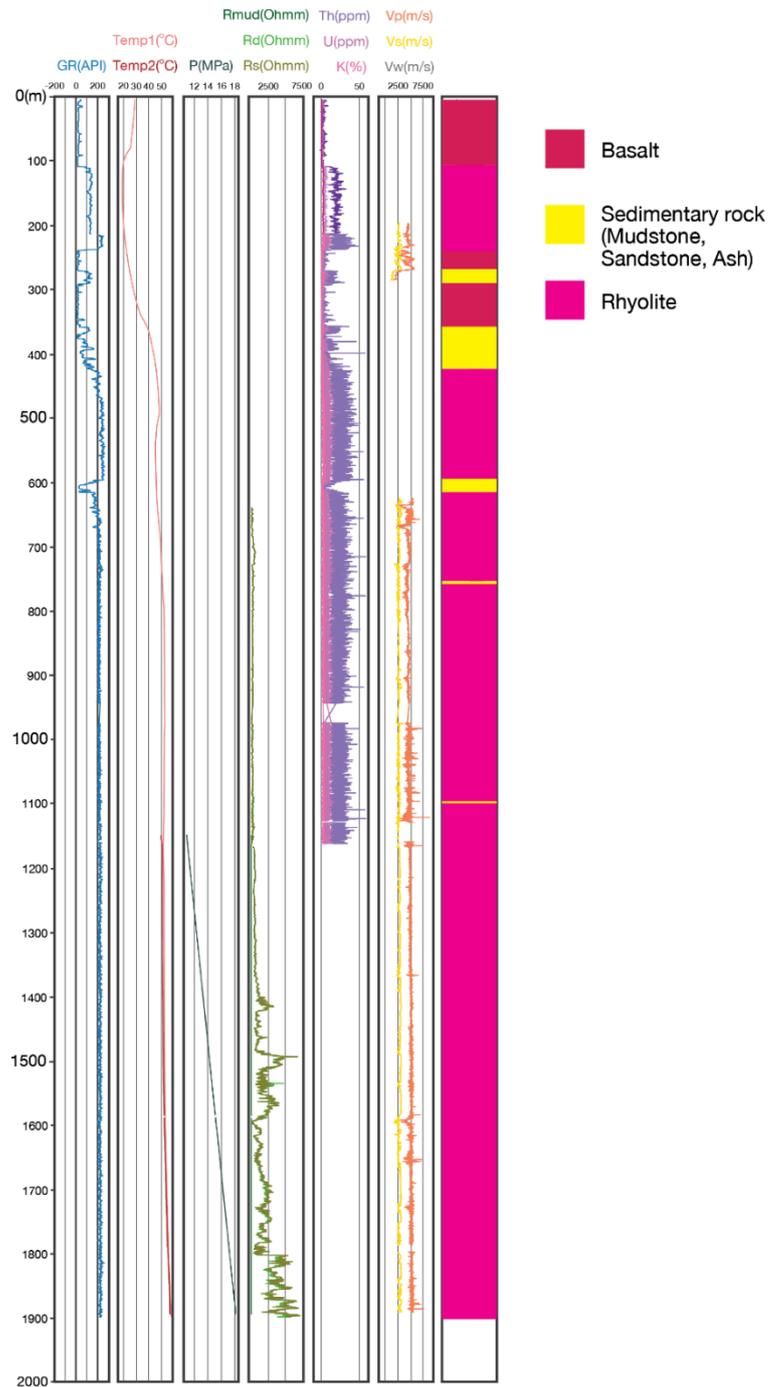


Figure 21 Well logging data of Kimberly with lithostratigraphic log. Kimberly well is drilled approximately 1,900 m with 0.1 m interval. GR: gamma ray, Temp: temperature, P: pressure, Rmud: resistivity in mud measurement, Rd: resistivity in deep measurement, Rs: resistivity in shallow measurement, Th: thorium, U: Uranium, K: Potassium, Vp: P-wave velocity, Vs: S-wave velocity, Vw: water wave velocity.

3.2 Snake River Plain

In Early Mesozoic Era, the oceanic Farallon plate, subducted underneath the North American plate. This subduction zone is called Farallon subduction zone and provides major tectonic features in west America as shown in Figure 22 (DeCourten, 2009). After that, the subducted Farallon plate failed and the fragment of Farallon plate beneath North American plate was melted in the mantle. This causes widespread volcanisms along the western of North American plate. As the next step, the fragments of Farallon plate were renamed into Juan de Fuca and Cocos plates (Lonsdale, 2005). During Early Cretaceous, worldwide plate motion was changed. Pacific plate began moving to the north away from North American plate. As a result, the strike-slip fault, San Andreas fault, had been created along the west margin of North American plate (Engebretson et al., 1984). Moreover, the Juan de Fuca subduction was rollback in this period. This subduction rollback caused North American plate moving westward over the Yellowstone hotspot where the fragment of Farallon plate was melted below North American plate and extension along west North American plate. Finally, Snake River Plain had been formed by these tectonic processes (Bedrosian and Feucht, 2014; Humphreys, 1995).

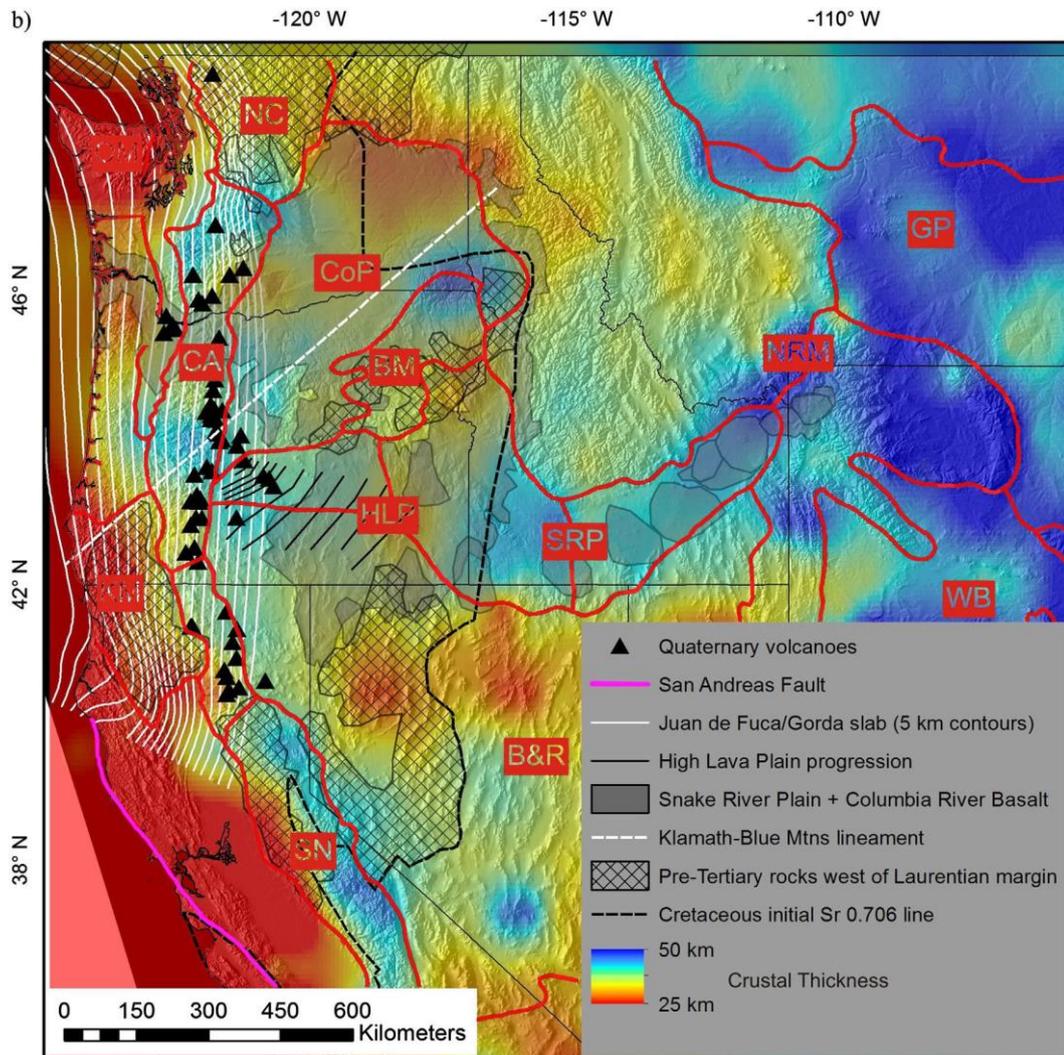


Figure 22 Crustal thickness map of western United State showing tectonic features and physiographic province: Basin and Range (B&R), Blue Mountains (BM), Cascade Arc (CA), Coast Ranges (CR), Colorado Plateau (CP), Columbia Plateau (CoP), Great Plains (GP), Great Valley (GV), High Lava Plains (HLP), Klamath Mountains (KM), North Cascades (NC), Northern Rocky Mountains (NRM), Olympic Mountains (OM), Sierra Nevadas (SN), Snake River Plain (SRP), Southern Rocky Mountains (SRM), and Wyoming Basin(WB) (Bedrosian and Feucht, 2014).

Six high-grade geothermal systems have been located in the western part of the US: the Great Basin, the Snake River Plain and margins, the Oregon Cascade Range, the Southern Rocky Mountains, the Salton Sea, and the Clear Lake Volcanic Field (Tester et al., 2006). In particular, 75% of the Snake River Plain, which is about 27,900-km², has an average temperature more than 200 °C at depth 4 km (McLing et al., 2016). The Snake River Plain is located in southern part of Idaho, extending about 6,400 km in an east-west direction. Furthermore, the Snake River Plain is a unique volcanic province which experienced bimodal volcanism during Quaternary. As a result, this area is covered by Quaternary basalt and rhyolite as shown in Figure 23 (Ellis et al., 2010). There are a large number of hot-springs. One of the famous hot-spring in this region includes the Yellowstone National Park, which attracts over 4 million visitors annually. (Brand and White, 2007; McLing et al., 2016).

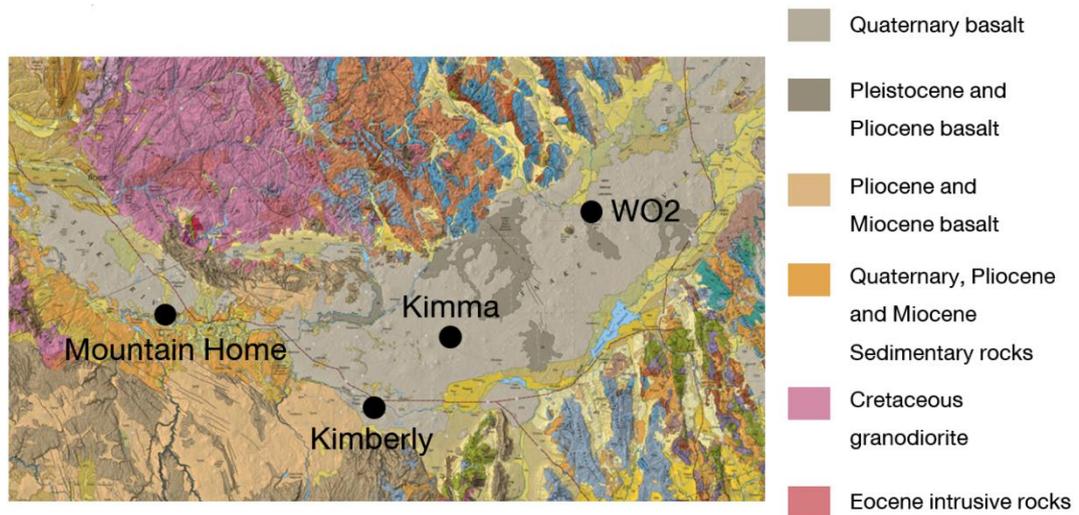


Figure 23 A geological map of Idaho, the USA, Snake River Plain with well location.

This area is covered by Quaternary basalt and rhyolite (Lewis et al., 2012).

The Snake River Plain is divided into two parts: Western Snake River Plain and Eastern Snake River Plain. The former is characterized by normal fault-bounded graben whereas the latter is related to the middle Miocene to recent volcanic activity with Yellowstone Hotspot, creating crustal downwarping, faulting, and caldera formation as shown in Figure 24 (Hughes et al., 1999; Neupane et al., 2014; Pierce and Morgan, 2009; Rodgers et al., 2002).

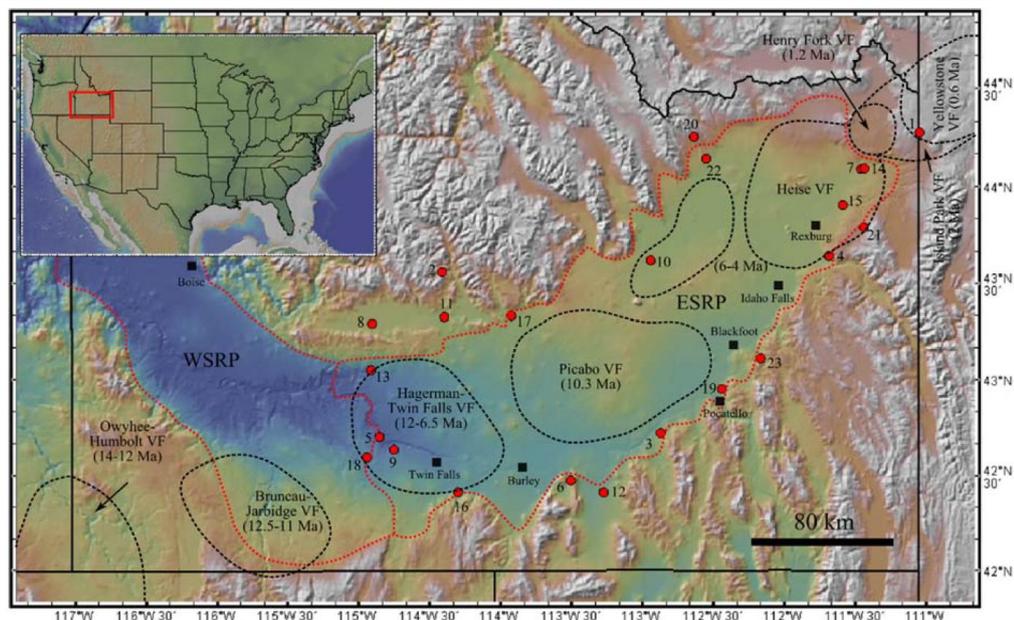


Figure 24 The Snake River Plain in south Idaho, USA. Red dash line represents the boundary between the Eastern Snake River Plain and the Western Snake River Plain. Black dash line refers to the area of Quaternary volcanic activity (Neupane et al., 2014).

3.2.1 Western Snake River Plain

Wood and Clemens (2002) proposes that the Western Snake River Plain is an intracontinental rift basin, which covers an area of 70 km wide and 300 km long. This basin is bounded by normal fault dipping toward the center of basin and filled with 2-3 km Neogene sedimentary rock (Figure 25). In comparison, the eastern part of

Snake River Plain in which there is no fault bounded is not rift basin but it is formed by the complex processes between extension and magmatism (Parsons et al., 1998). Therefore, Western Snake River Plain is more simplified than Eastern Snake River Plain in terms of both tectonic and magmatic activities. Although the structure of Western Snake River Plain is similar to typical intracontinental rift basins like East African Rift Valley, Baikal Rift, and Rio Grande Rift with these including half grabens and full grabens, the magmatism in Western Snake River Plain is different. This area has been affected by hotspot beneath Snake River Plain (Wood and Clemens, 2002).

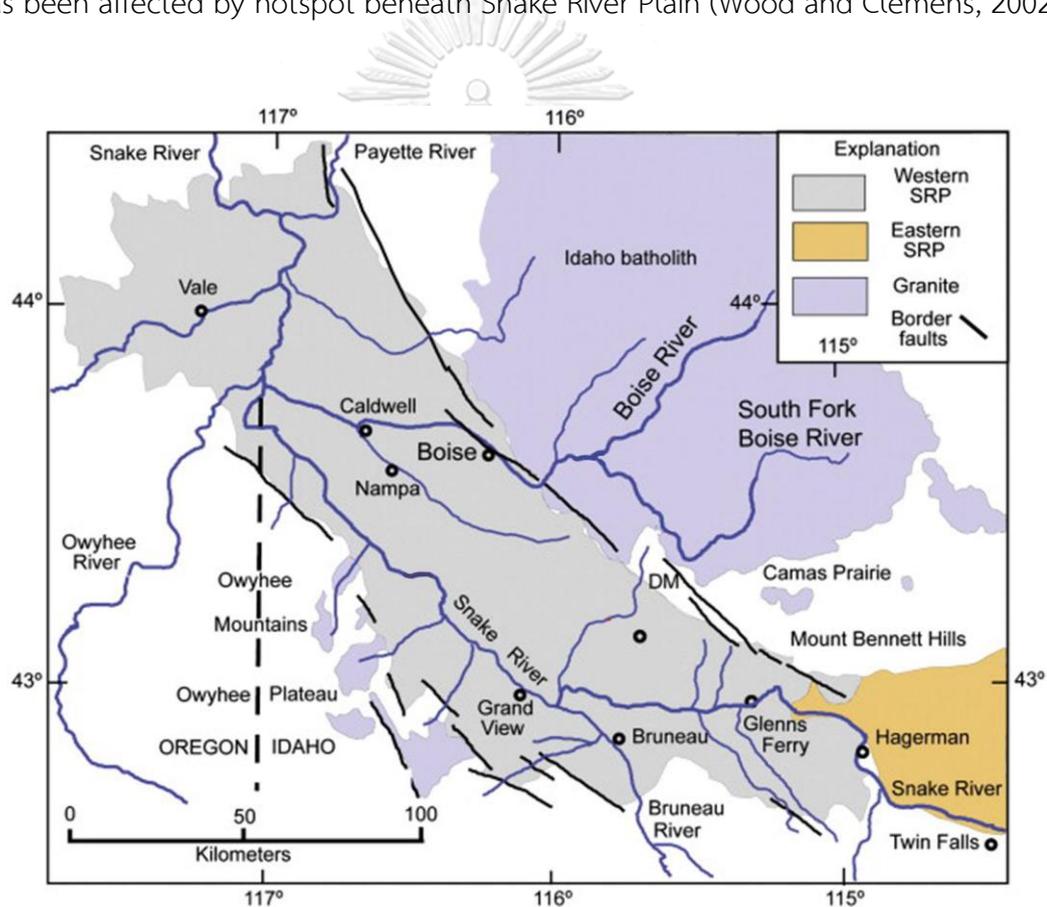


Figure 25 The intracontinental rift basin Western Snake River Plain. Black lines are border normal fault dipping toward the center of Western Snake River Plain (Shervais and Vetter, 2009).

In the beginning, when Western Snake River Plain basin was extending, lacustrine environment was a major system in this basin and then it changed into fluvial system in Late Quaternary. Moreover, there were rhyolitic volcanic activities during pre-rift and basaltic volcanic activities in post-rift. Therefore, this area consisted of sedimentary rocks from lacustrine and fluvial environments such as sandstone, mudstone, and conglomerate and igneous rocks like rhyolite, basalt, and tuff. The cross-section in Western Snake River Plain is shown in Figure 26.

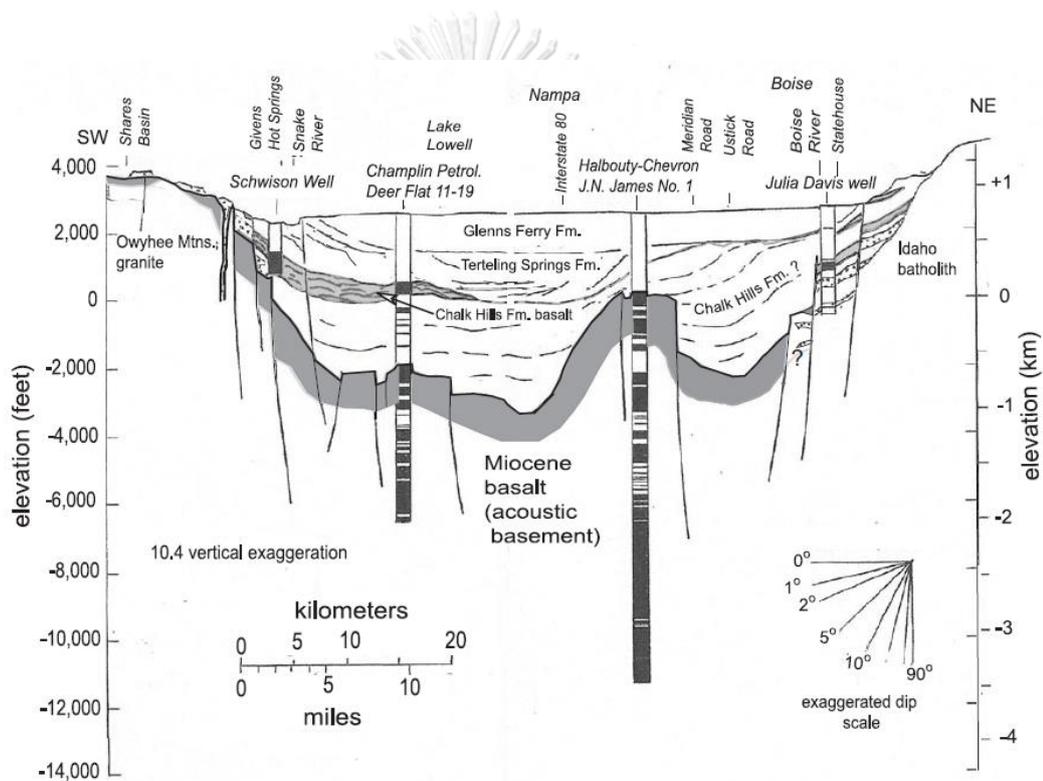


Figure 26 Cross section of Western Snake River Plain. The dark pattern shows basalt in this area (Wood and Clemens, 2002).

Several geologists interest in Western Snake River Plain because of large groundwater resource and geothermal energy. Due to semiarid of southwest Idaho, groundwater resource is necessary in term of economy and society. A lot of data from failed petroleum wells help geologists improve accuracy of groundwater models (Whitehead, 1992). By doing this, aquifer distribution is more clearly understood and groundwater management is improved. However, geothermal resource in Western Snake River Plain is less interested than Eastern Snake River Plain because Eastern Snake River Plain has various magmatism than this area.

Before formed into Western Snake River Plain, this area was directly affected by rhyolitic magma and Steens basalt during 17-13 Ma. Then, rhyolitic magma erupted along fault margin while the basin was developing, 12-11 Ma. Moreover, convincing evidence, 2 km thick of continuous rhyolite sequence, from deep well in the center of the western plain shows that Bruneau-Jarbidge rhyolite eruptive center coincides with formed Western Snake River Plain basin. After that normal faults rapidly developed, followed by extending over 2 km offset from 11 Ma to 9 Ma and then movement rates have been low (<0.01 mm/year).

Depositional environment of this basin is fluvio-lacustrine interbedded with volcanic rock in some periodes. The sedimentary rocks of this basin are represented by Chalk Hills Formation determined by Perkins et al. (1998) comprising claystone, arkosic sandstone, tuff, and volcanic rock as shown in Figure 27. Due to erosional surface founded in Chalk Hills Formation, Wood and Clemens (2002) interpreted that there was regressive of lake water in some period between 6 and 4 Ma and then water level rebounded with the depositing of transgressive sequence over the erosional surface. Moreover, oolites found in this area indicates a rise of alkaline in lake water and then disappearing of oolite indicates declining of alkaline because of connecting to river and sea. Tassell et al. (2001) suggested that this lake connected to the Snake River, the Columbia-Salmon River, and the sea between 3.8 and 2 Ma.

Finally, the environment gradually changed to fluvial environment from southeast to northwest. As a result, the upper section of Chalk Hills Formation is deltaic sandstones and fluvial deposits.

After filled with lake sediments, Western Snake River Plain had effect from basalt around 3.2 Ma. This basalt overlaying on sedimentary rocks expressed source from shield volcanoes and fissure eruption. In Late Quaternary, the southeastern boundary fault system was reactivated by the present tectonic stress. This caused fault movement and orientation.



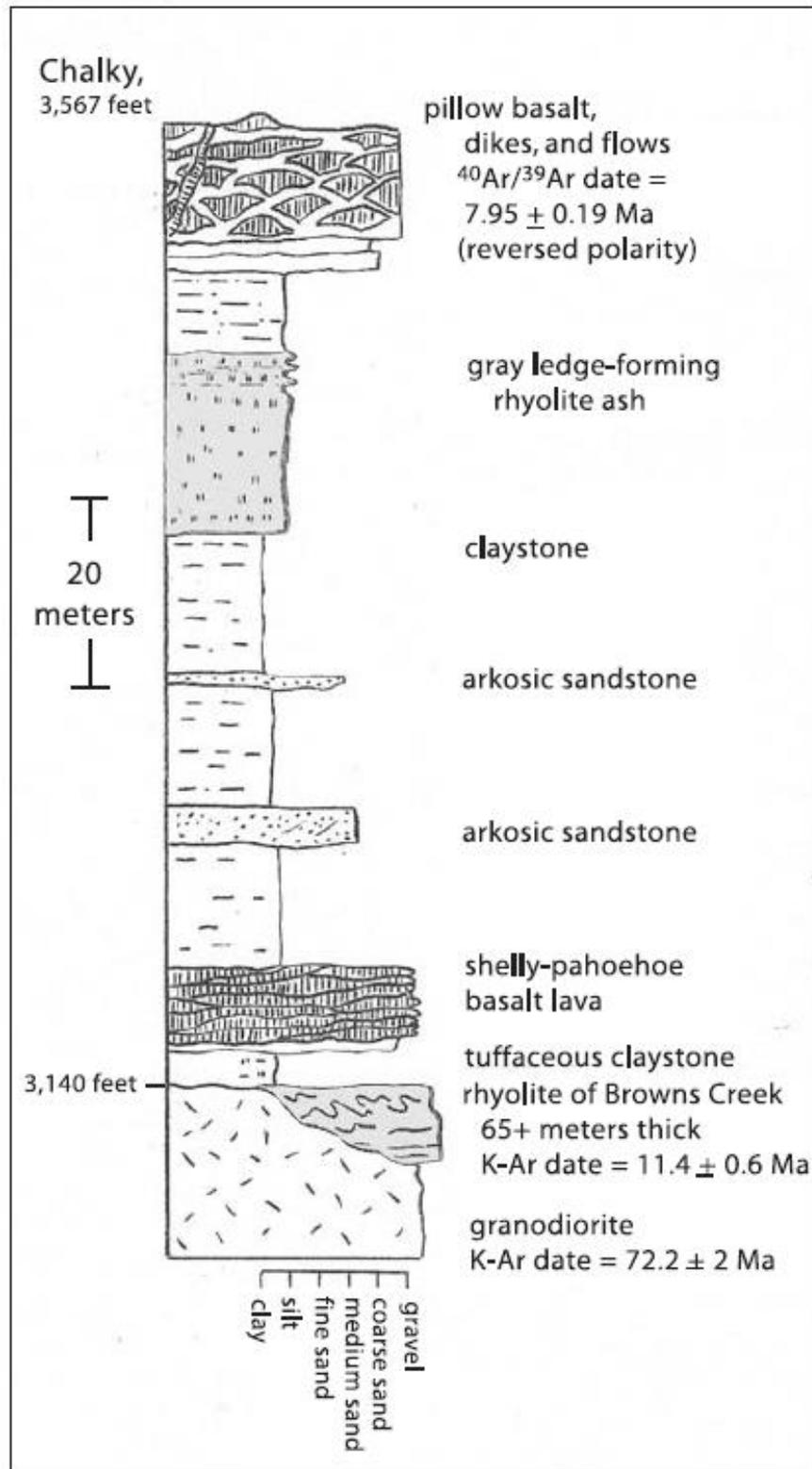


Figure 27 Stratigraphic column of Chalk Hills Formation represented rock in the Western Snake River Plain (Wood and Clemens, 2002).

3.2.2 The Eastern Snake River Plain

Eastern Snake River Plain is relatively low land area compared with the adjacent areas like Basin and Range, Blue Mountains, and Northern Rocky Mountains as shown in Figure 28. Structure and geomorphology obviously show whether there are different tectonic processes with Eastern Snake River Plain and Basin and Range where are dominated by normal faults and high elevation. Basin and Range has had extremely deformation more than Eastern Snake River Plain. Moreover, when compared with the Western Snake River Plain, this area has rarely normal faults and basaltic rift zones because it has different extension processes and volcanic activities. The cross section of Eastern Snake River Plain is shown as Figure 29. The first stage of Eastern Snake River Plain began around 16 Ma with basin extension and then it has subsided through the Neogene.

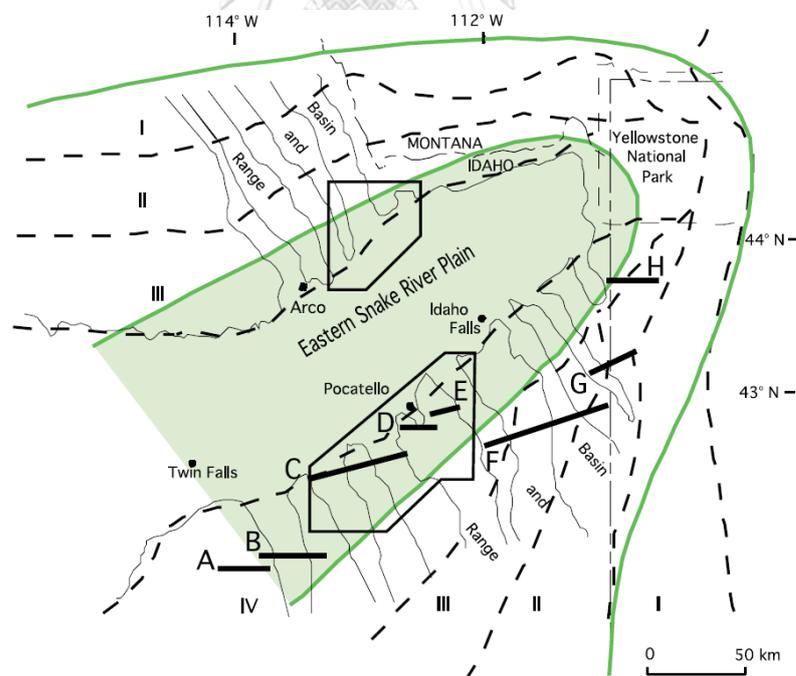


Figure 28 Eastern Snake River Plain bounded by Basin and Range and seismic parabola (green line) (Rodgers et al., 2002).

Eastern Snake River Plain is adjacent to Basin and Range and bounded with Quaternary fault zone called seismic parabola (Figure 28) determined by Pierce and Morgan (1992). Moreover, founded structures in Eastern Snake River Plain indicate whether Eastern Snake River Plain and Basin and Range are formed by similar processes but Eastern Snake River Plain has been affected by a normal fault and dike injection. The dikes in Pliocene-Quaternary are mainly caused by extension in this area (Kuntz et al., 1992).

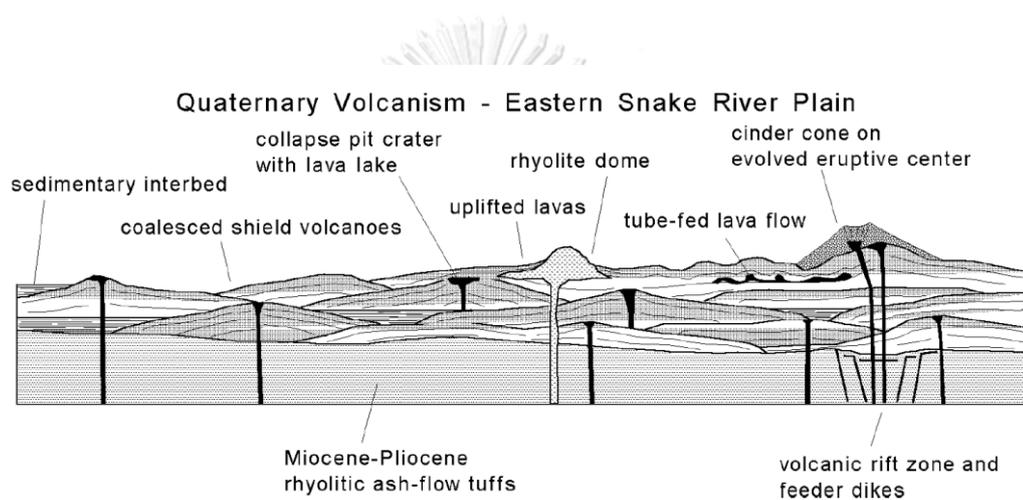


Figure 29 The typical cross section of Eastern Snake River Plain. There are more volcanic eruptions in this area than Western Snake River Plain.

Rodgers et al. (2002) summarized data from Pierce and Morgan (1992) and McQuarrie and Rodgers (1998) and proposed model to explain how Eastern Snake River Plain form as shown in Figure 30.

Stage 1 during 16 - 11 Ma, the proto-Eastern Snake River Plain was formed with slow extension and subsidence rate, followed by sedimentation rates being about 50-80 m/m.y. The proto-Eastern Snake River Plain basin was bounded by mountains or highlands and half graben was developing during this stage. On this stage, there was rarely volcanic activities so merely some basalts and tuffs were found. Furthermore, the proto-Eastern Snake River Plain basin was filled with thin

sequences of carbonate rocks and fine grained clastic rocks as an environment in this stage was interpreted to lacustrine and fluvio-lacustrine.

Stage 2 during 11-4 Ma, in this stage, Eastern Snake River Plain rotated to the current position on the Yellowstone Plateau. Moreover, the extension rate in Eastern Snake River Plain was increased before it decreased in stage 3. This resulted high rate subsidence and thick sedimentary rocks deposited with sedimentary rates of 100-300 m/m.y. Furthermore, the upper crust was extended by normal fault and the lower crust was extended and thinned by heat flow and dike injection. Due to thin crust, mafic magma could intrude into this area and this dikes and sills encouraged Eastern Snake River Plain to subside faster. After that, from 8.5 to 6.0 Ma, this area directly was affected by silicic magmatism. It caused widely highland topography on this area.

Stage 3 during 4-0 Ma, the extension direction changed because of rotation in stage 3. In this stage, Eastern Snake River Plain has slightly subsided and Quaternary faults have developed. Moreover, there are various volcanic activity and hot spring around Yellowstone Plateau.

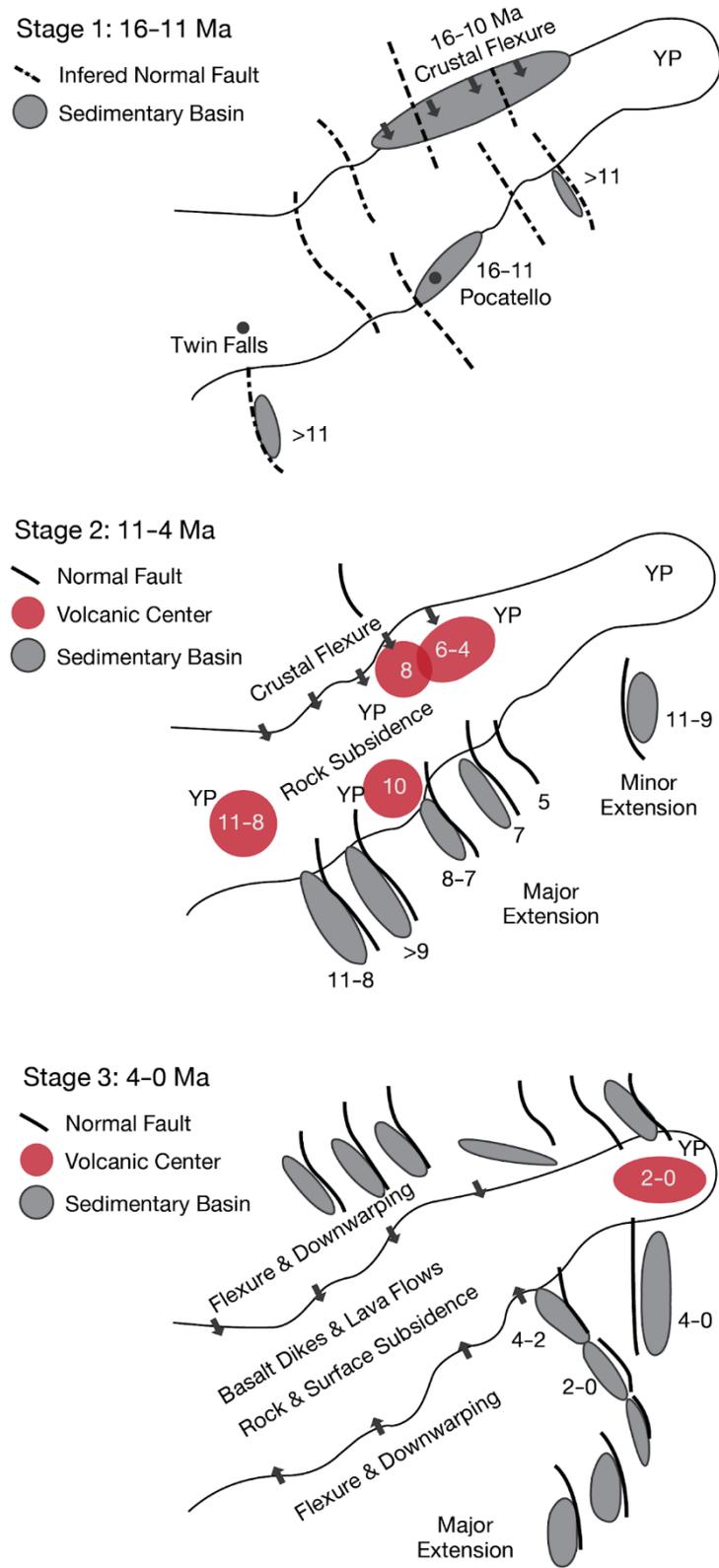


Figure 30 Evolution of Eastern Snake River Plain in stage 1 to 4 proposed by Rodgers et al. (2002).

Chapter 4 Methodology

4.1 Overview of Workflow

In this study, procedures are divided into main three steps: data preparation, models development, and models comparison. In Step 1, data are prepared for the models in data preparation or pre-processing. The data are cleaned and divided into training data, validation data, and test data (Figure 31). Details of data preparation are explained in section 4.2

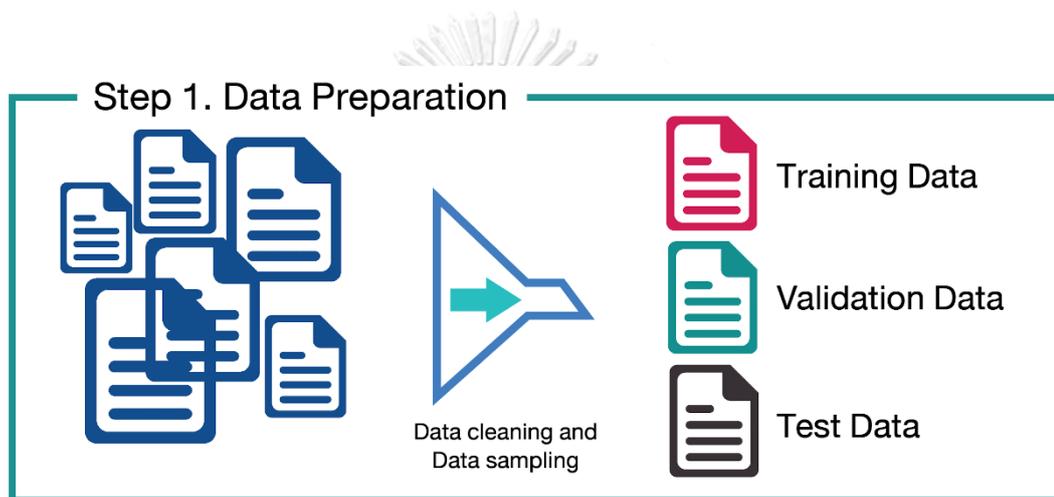


Figure 31 Data preparation step. The data are cleaned and split into training, validation, and test data.

In Step 2, training data are used to train and build machine learning algorithms. The models are trained to know the characteristic of each rock by training data. After that validation data are used to evaluate preliminary results and some parameters of the models are changed. The models are optimized by validation data iteratively until the optimal parameters are found. This process is called hyperparameter tuning and the details of hyperparameter tuning are presented in section 4.3.2. As the next step, the optimal models are evaluated by test data to give an accuracy score of each algorithm. Figure 32 shows the diagram of the models development step. Details of models development are explained in section 4.3.

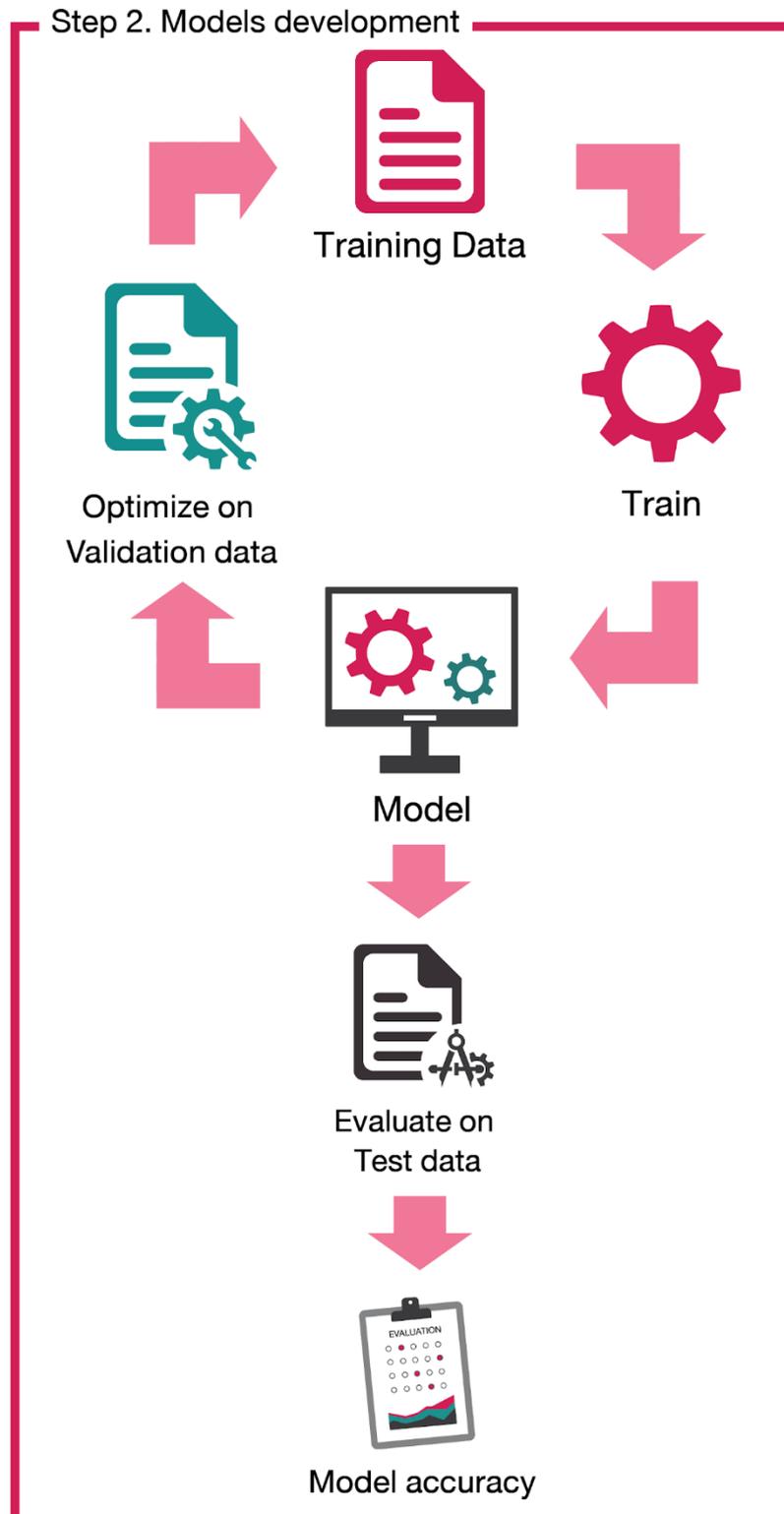


Figure 32 Models development step. The models are trained by training data and optimized by validation data. After that, they are evaluated by test data.

As six algorithms are employed in this study, Step 3, models development, is iteratively done with six machine learning algorithms. Due to optimization, the models are led to overfit with validation data. This reduces the generalization of our algorithms for other data. Hence, the same data are prepared into training, validation, and test data with five different data sampling. Then each model is trained, optimized, and tested five times with different data sampling five times and the accuracy score of each algorithm is an average accuracy from five tests. By doing this, the variance from test data is reduced. This is called Monte Carlo cross-validation (Xu and Liang, 2001). Finally, the performance of each algorithm is compared in step 3, the models comparison, as shown in Figure 33. Details of models comparison are shown in section 4.4.

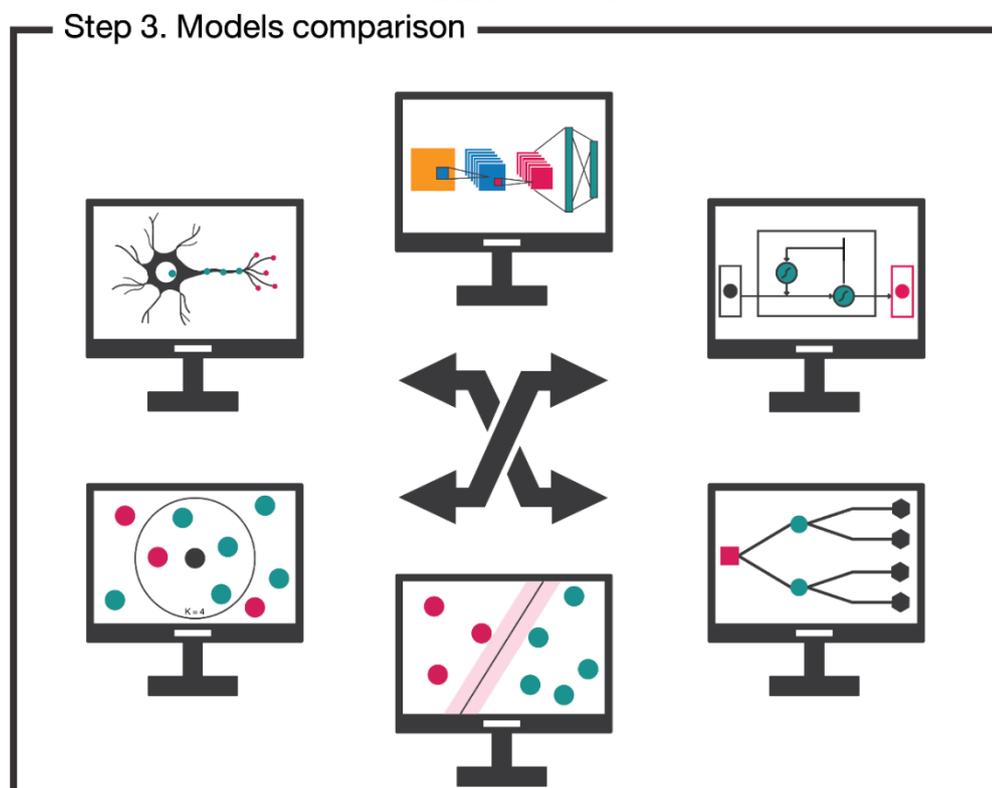


Figure 33 Models comparison step. The performance of six models which are optimized by validation data is compared in this step.

4.2 Data preparation

Well logging data are correlated with lithology from well reports. After every data point is labeled, the process of cleaning data begins. Data points which have less than three features and also negative data points are removed. A problem with this dataset is missing information from the original dataset. This can be due to the error for well log readings or the lack of particular well logging surveys in some wells. There are various techniques to tackle this problem. Generally, statistical approaches such as replacing missing value with the mode or median are used for data which have small missing values. For data which have a lot of missing values, missing imputation method is proposed to fill the missing values. For example, Tsai et al. (2018) proposed Class Center based Missing Value Imputation for missing values imputation. Class Center based Missing Value Imputation calculates missing values from surrounding data point as shown in Figure 34. Another example is Lopes and Jorge (2017) that missing values are filled by machine learning algorithms. A study by Lopes and Jorge (2017) shows that ANN gives the best performance in replacing missing values, followed by providing lower than 5% of error. However, Lopes and Jorge (2017) requires about 600,000 data points for machine learning algorithms. As this study does not have much data to do that, since the missing data points are replaced with -999.25 instead. As Shervais (2014a) uses -999.25 for absent value, this study uses that number too. This study aims that the model will recognize -999.25 is a missing value. After pre-processing is finished, the data are prepared for each experiment in the next process.

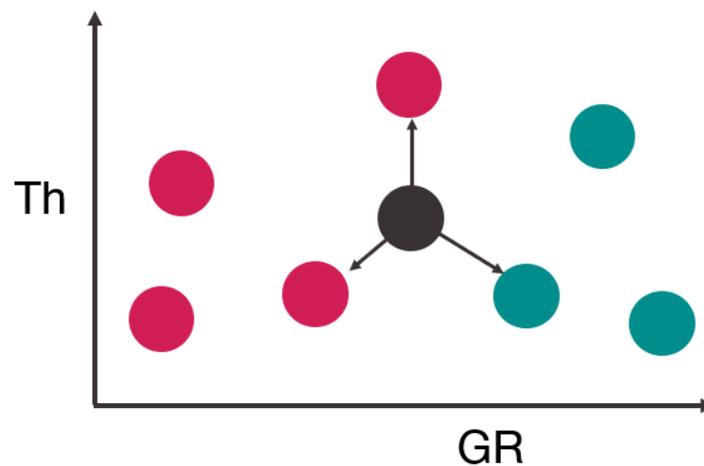


Figure 34 The conceptual model of Class Center based Missing Value Imputation. It calculates missing values for surrounding data points similar to K-nn (Tsai et al., 2018).

The data are divided into training, validation, and test subsets depending on each experiment by stratified sampling. Stratified sampling is a data sampling method which data are divided into subset while preserving class distribution (Figure 35). The training dataset is used to train the models. The validation dataset is applied to determine the optimal tuning parameters. The test dataset is employed in order to gain the classification accuracy and confusion matrix. The training, validation, and test data are split differently five times in each experiment to reduce variance of data.



Figure 35 A diagram shows stratified sampling where the data are split while class distribution remains constant.

4.3 Models development

Generally, machine learning algorithms are trained, optimized, and tested with training, validation, and test data, respectively. However, there are many possible ways to divide the training, validation, and test data. For example, training, validation, and test data are randomly selected from the same well (Xie et al., 2018). One well is determined as test data and other wells are determined as training and validation data (Hall and Hall, 2017). In this study, how to divide the data into training, validation, and test data are called as experiments or experimental design. As different experimental design might give disparate results in machine learning algorithms, this study does both experiments: single-well and multiple-well test.

4.3.1 Experimental design

Machine learning algorithms are employed under two experiments to simulate real-world application for this study: single-well and multiple-well test. In the single-well test, well logging data from each well are randomly split into training (70%), validation (10%), and test (20%) subsets, and then used to train and test with all six models as shown in Figure 36. For this experiment, well logging data and classes of each well are kept independent from those of other wells.

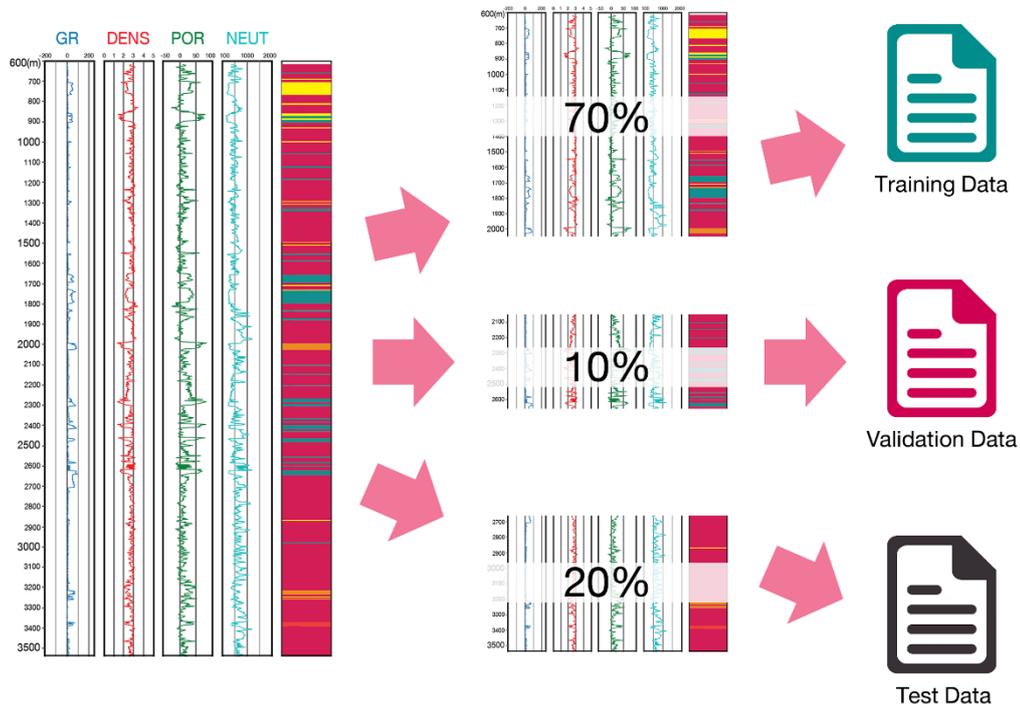


Figure 36 The single-well test. The training (70%), validation (10%), and test (20%) data are randomly split from the data in each well.

The combined well test uses a large amount of data as a training subset, combining data from three wells and 70% of the data from the remaining well (Figure 37). Then the remaining data are separated into validation (10%) and test (20%) subsets. As this scenario is to combine well logging data from four wells, the data have to be classified with the same classification in every well. In this scenario, lithology is reclassified into five classes: basalt, sedimentary rock, tuff, vitrophyre, and rhyolite. This is because sedimentary rocks in WO2 are classified as claystone, sandstone, siltstone while the other wells grouped as sedimentary rock. Features are prepared into 15 features while missing values are filled with -999.25. As some features were not collected in every well, their values are -999.25 in some wells. For example, the values of temperature logs are -999.25 in every data point in WO2 well.



Figure 37 The multiple-well test. Three wells and 70% of the fourth well are assigned as training data while the rest of the fourth well is assigned as validation (10%) and test (20%) data.

4.3.2 Hyperparameter tuning

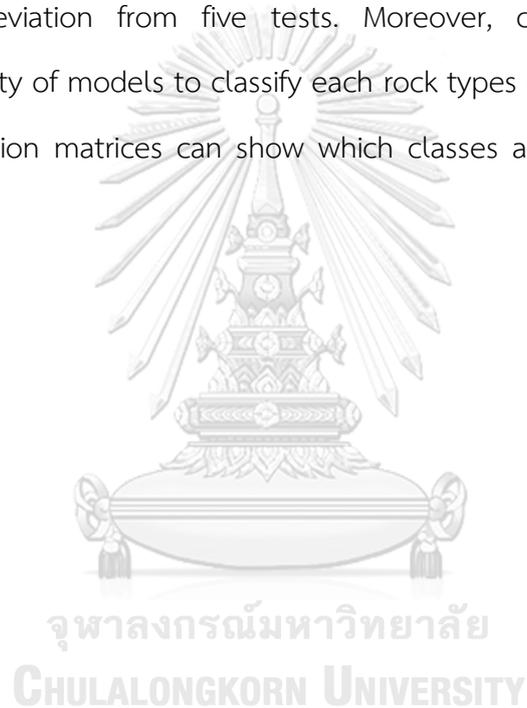
Hyperparameter or tuning parameter is essential to improving the performance of machine learning algorithms. A robust parameter selection process (tuning) is a process which ranks the accuracy of each model with different parameters to obtain the optimal parameters for each algorithm. Tuning parameter ranges for this study are exhibited in Table 3. The optimal parameter ranges are determined with the help of the validation dataset.

Table 3 A range of hyperparameter tuned for each algorithm.

Model	Tuning Parameter	Search range
SVM	Penalty parameter of the error term (C)	0.1 - 1000
	Kernel coefficient for 'RBF' (γ)	0.00001 - 0.01
K-nn	Number of K	1-10
XGB	Learning rate	0.01 - 0.3
	The minimum number of samples required at a leaf node (min child weight)	0.1 - 100
	Maximum depth of the individual tree (max depth)	3 - 10
ANN	Learning rate	0.00001 - 0.01
	Dropout	0.05 - 0.25
	Number of node	8 - 64
CNN	Learning rate	0.00001 - 0.01
	Dropout	0.05 - 0.25
	Filter	8 - 64
GRU	Learning rate	0.00001 - 0.01
	Dropout	0.05 - 0.25
	Unit	8 - 64

4.4 Models comparison

The performance of six models evaluated by test data is compared in this step. As every rock types (or classes) is important equally, the accuracy score is used to evaluate the performance of each model in this study. To prevent overfitting, the data are divided into training, validation, and test five times with different data sampling and the models are trained and tested five times in each experiment. Hence, the accuracy score in each experiment of this study is an average accuracy with standard deviation from five tests. Moreover, confusion matrices which represent the ability of models to classify each rock types or classes are generated in this study. Confusion matrices can show which classes are classified correctly and which classes not.



Chapter 5 Results

5.1 Model Performance

This study makes an attempt to design experiments in real-world environment so experiments are divided into two experiments. The optimal parameters are obtained by ranking models with different parameters in validation dataset and then they are applied to the models for testing on the test dataset. Each experiment is iterated five times to reduce variance of data. The accuracies of each model are given by mean of accuracy from five times test and the average accuracies of each model in each of experiments are shown in Table 4. Preliminary results show that XGB is the best model for lithological classification by using well-logging data whereas neural network models, ANN, CNN, and GRU, are the worst performing model. Figure 38 shows the comparison results between well report and the results from each model.

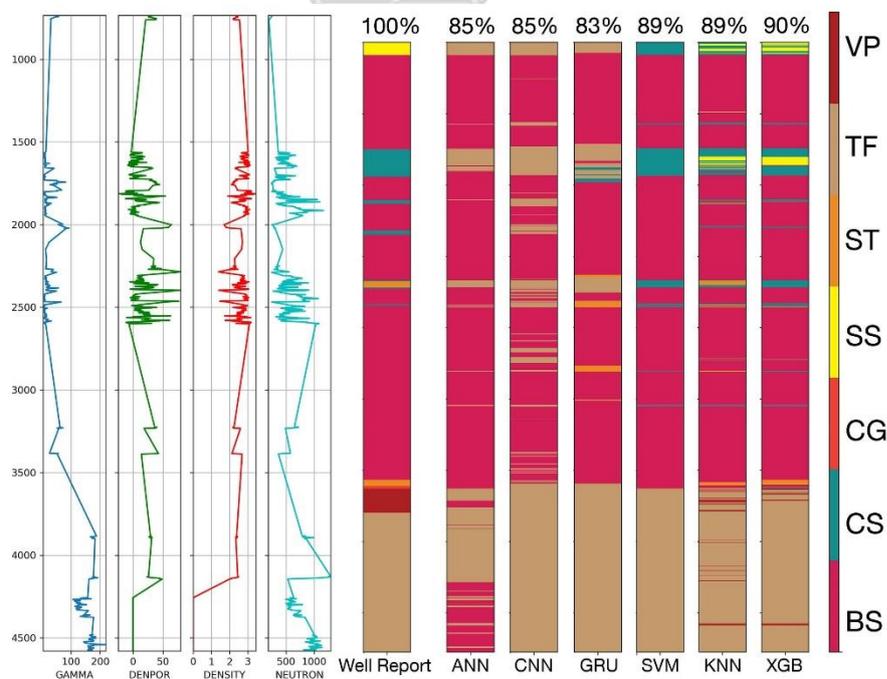


Figure 38 The comparison results between well report and the models. XGB gives the highest accuracy score at 90%. SVM and K-nn provide comparable results at 89% while ANN, CNN, and GRU exhibit the lowest accuracy.

Table 4 Classification accuracies of each algorithm for single- and multiple-well tests.

Model	Accuracy	
	Single-well test	Multiple-well test
SVM	91% ± 8	84% ± 21
KNN	88% ± 11	86% ± 15
XGB	91% ± 9	87% ± 13
ANN	88% ± 13	78% ± 21
CNN	84% ± 14	77% ± 21
GRU	86% ± 14	78% ± 19

5.1.1 Single-well test

Well-log data of each well are randomly split into training (70%), validation (10%) and test (20%) subsets and used to train and test with all six models. Results suggest that XGB and SVM give the comparable accuracy of rock classification (91%) as shown in Figure 39. K-nn show prediction accuracy at $88\% \pm 11$ while ANN, CNN, and GRU, display comparable accuracy at $88\% \pm 13$, $84\% \pm 14$, and $86\% \pm 14$, respectively. Table 5 presents the representative accuracy of each rock in the single-well test. The representative accuracy is averaged from all wells and selected from one out of five folds in Monte Carlo cross-validation. In general, igneous rocks such as basalt, tuff, and rhyolite have high accuracy (more than 80%) in every model. However, some models often misclassify particular rock types such as sedimentary rocks and vitrophyre, resulting in 0% average accuracy.

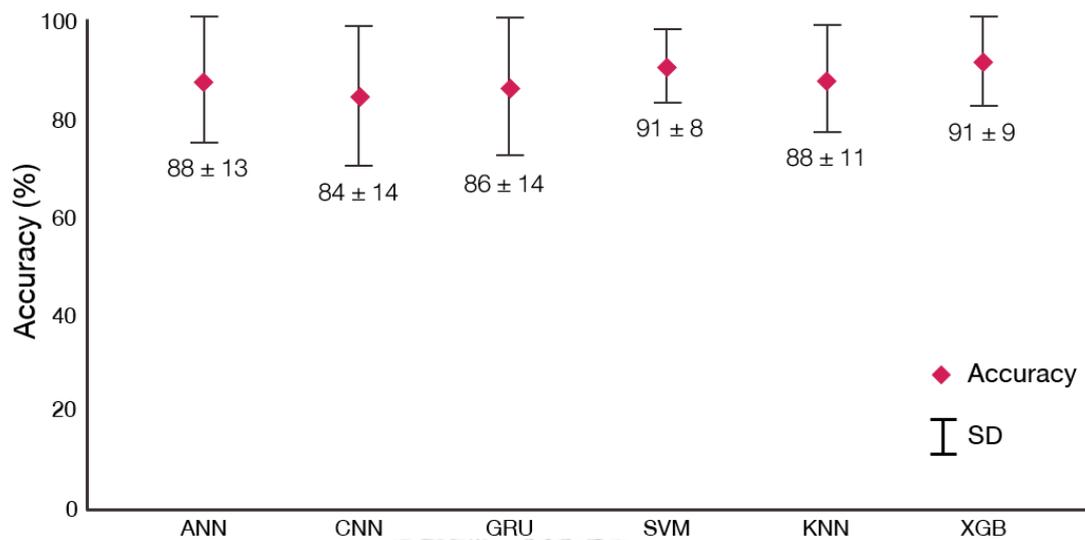


Figure 39 Average accuracy scores over six machine learning algorithms in single-well test. Red spots represent min and max accuracies of each algorithm before average. Error bars represent standard deviation.

Table 5 The representative accuracy of each rock type from the single-well test. The representative accuracy is averaged from all wells and selected from one out of five folds in Monte Carlo cross-validation.

Rock types	SVM (%)	KNN (%)	XGB (%)	ANN (%)	CNN (%)	GRU (%)
Basalt	94	93	93	73	73	72
Claystone	75	50	54	0	0	0
Conglomerate	0	25	0	0	0	0
Sandstone	0	39	61	0	0	0
Siltstone	0	56	41	0	0	0
Vitrophyre	0	28	16	0	0	0
Tuff	100	97	99	100	100	100
Sedimentary rock	20	20	52	5	5	12
Rhyolite	87	53	75	88	95	96

Remark: 0% means the model misclassify this rock completely and 100% means the model classify this rock perfectly.

5.1.2 Multiple-well test

Multiple-well test uses the largest amount of data as a training set, which combines data from three wells and 70% of data from the remaining well. To produce comprehensive results, the test set is also permuted between wells until every well has become test set. However, sedimentary rocks are grouped as sedimentary rock class in this experiment. There are five classes: basalt, sedimentary rock, tuff, vitrophyre, and rhyolite with 15 features in this scenario. Results are similar to the single-well test as shown in Figure 40. XGB remains the best algorithm for lithological classification, followed by $87\% \pm 13$ of accuracy. The next-best performance models are K-nn and SVM, showing $86\% \pm 15$ and $84\% \pm 21$, respectively. CNN gives the lowest accuracy among the neural network models at $77\% \pm 21$ whereas ANN and GRU give a commensurate performance at $78\% \pm 21$ and $78\% \pm 19$, respectively. Table 6 shows the representative accuracy of each rock in multiple-well test. The accuracy of rhyolite classified by SVM, K-nn, and XGB is reduced from single-well test whereas the accuracy of tuff exhibited by deep learning algorithms is declined from single-well test.

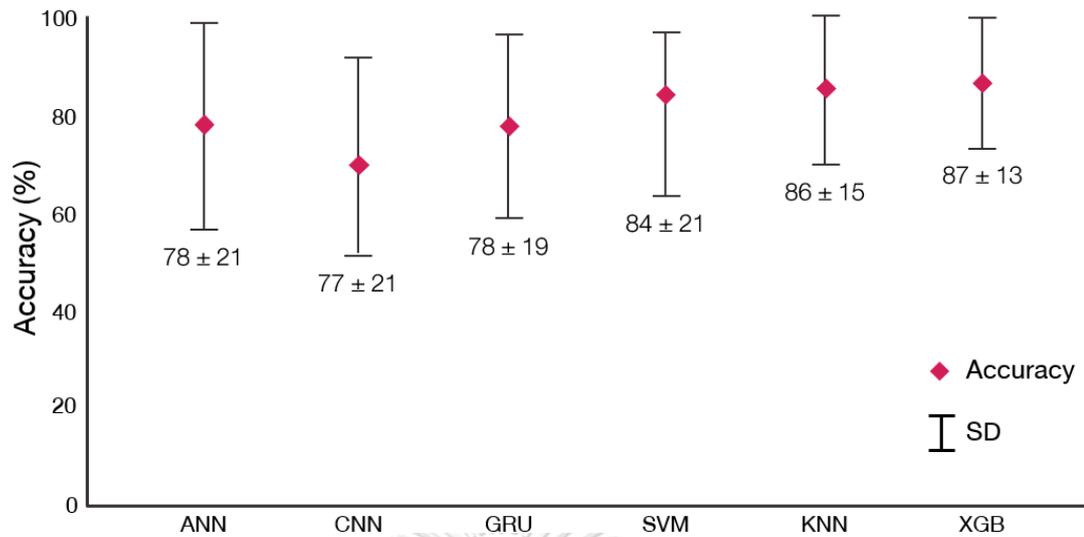


Figure 40 Average Accuracy scores over six machine learning algorithms in multiple-well test. Red spots represent min and max accuracies of each algorithm before average. Error bars represent standard deviation.

Table 6 The representative accuracy of each rock type from the multiple-well test. The representative accuracy is averaged from all wells and selected from one out of five folds in Monte Carlo cross-validation.

Rock types	SVM (%)	KNN (%)	XGB (%)	ANN (%)	CNN (%)	GRU (%)
Basalt	100	100	100	85	75	98
Sedimentary rock	31	47	41	0	0	13
Tuff	82	86	99	0	0	0
Vitrophyre	0	19	0	0	0	0
Rhyolite	58	59	63	91	98	89

Remark: 0% means the model misclassify this rock completely and 100% means the model classify this rock perfectly.

5.2 The effect of tuning parameter

Machine learning algorithm is an automated system that automatically builds a model from the data without requiring time consuming and human involvement. However, there is no perfect model which can represent any datasets in the world. The models should be adjusted by changing some parameters (tuning parameters) to specify the models with one dataset. Consequently, tuning parameters or hyperparameters are essential to improve the efficiency of machine learning algorithms. The models are trained and tested five times to discourage the effect of random seed and to encouraged generalization of our algorithms. The optimal and search range of hyperparameters are exhibited in Table 5. The results show that the models are robust with the close range of hyperparameters in single- and multiple-well tests. The optimal hyperparameters are given by ranking each model on validation dataset since they do not always provide the highest accuracy on test data.

Table 7 The optimal range of hyperparameter tuned for each algorithm in single- and multiple-well tests.

Model	Tuning parameters	Search range	Optimal range
SVM	Penalty parameter of the error term (C)	0.1 - 1000	0.1 - 100
	Kernel coefficient for 'RBF' (γ)	0.00001 - 0.01	0.00001 - 0.0001
K-nn	Number of K	1-10	6 - 10
XGB	Learning rate	0.01 - 0.3	0.01 - 0.1
	The minimum number of samples required at a leaf node (min child weight)	0.1 - 100	60 - 100
	Maximum depth of the individual tree (max depth)	3 - 10	3 - 6
ANN	Learning rate	0.00001 - 0.01	0.00001 - 0.001
	Dropout	0.05 - 0.25	> 0.25
	Number of node	8 - 64	16 - 64
CNN	Learning rate	0.00001 - 0.01	0.00001 - 0.001
	Dropout	0.05 - 0.25	> 0.25
	Filter	8 - 64	16 - 64
GRU	Learning rate	0.00001 - 0.01	0.00001 - 0.001
	Dropout	0.05 - 0.25	> 0.25
	Unit	8 - 64	16 - 64

Regarding approaches, there are two tuning parameters for SVM: C and γ . C controls the number of misclassified data in SVM with soft margin. Results show that C should be a small number between 0.1 and 100 (Figure 53 in Appendix A). This is because the generalization of SVM should be increased while misclassified data are restrained. Likewise, γ reflecting the radius of the decision boundary should be a small number too. Even though increased γ in single-well test rise accuracy of validation dataset, it reduces the accuracy of test dataset. Thus, this study suggests that γ should be 0.00001 to 0.0001.

For KNN, K is tuned. It depends on coherence and incoherence between training and test data. If training data are clean and consistent with test data, K parameter should be a small number (2 - 4). In comparison, if the training data consist of a lot of noises and are not consistent with the test data, K should be a big number (8 - 10). There are plenty of rock types which have overlapping feature ranges such as basalt and rhyolite or sandstone and conglomerate in this dataset. This leads K-nn to misclassify the actual rock types to their counterparts. Moreover, replacing missing value with -999.25 might extend the distance between the test data point and suitable data points. Consequently, the optimal K should be a big number for this dataset as shown in Figure 54 (Appendix A). Results from this study suggest that K should be an even number from 6 to 10 in single- and multiple-well tests.

Learning rate, min child weight, and max depth are tuned for XGB. Learning rate reflects how much model changes when new data are being added. If the learning rate is too small, the model will be trapped in local minima. In contrast, too much learning rate encourages model to pass the optimal. This study suggests that the optimal range of learning rate is 0.01 to 0.1. Min child weight refers to the number of instances in each node and max depth controls the maximum depth of the decision tree. The optimal range of min child weight is 60 to 100. However, the

effect of min child weight to model for validation and test dataset in multiple-well test is slightly different, followed by varying 2% of accuracy as shown in Figure 41. The efficiency of XGB is reliance on learning rate and min child weight more than max depth. The optimal range of max depth should be 3 to 6.

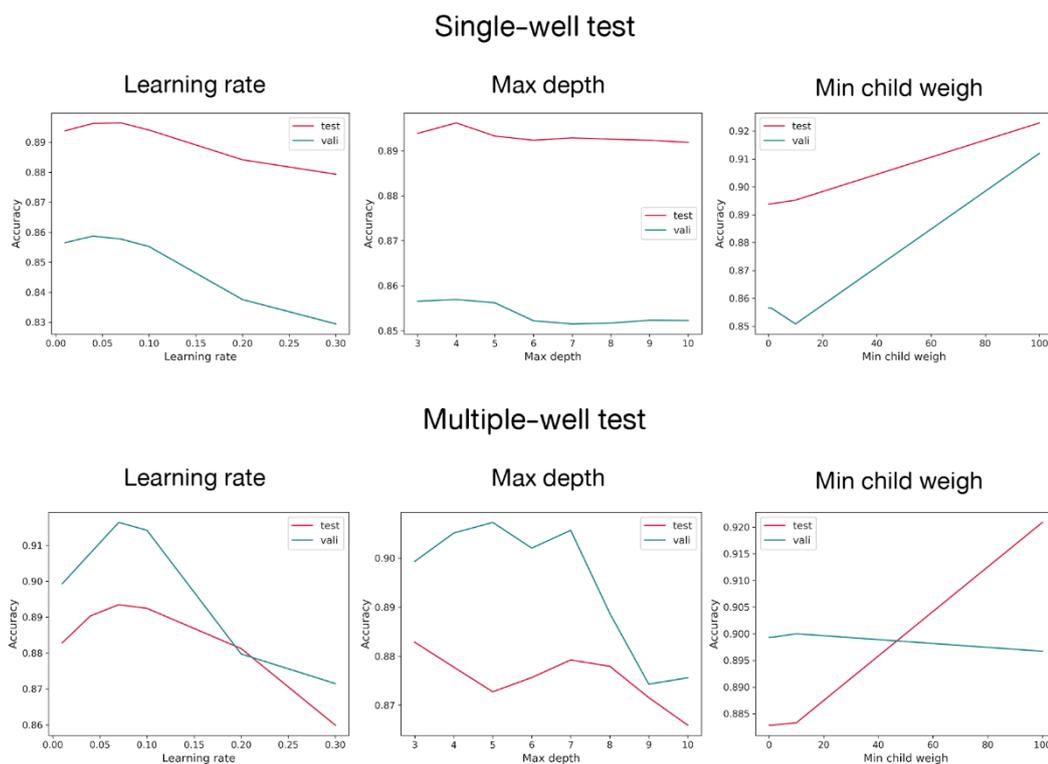


Figure 41 The line graphs show the effect of tuning parameter to accuracy in XGB in two experiments. Red line represents the accuracy of test data and validation data for green line.

ANN, CNN, and GRU apply the same strategy to classify data and the tuning parameters are almost the same. Hence, ANN, CNN, and GRU are grouped as neural network models. Neural network models exhibit lower prediction accuracy than other models because they require balanced data to train models. There are three tuning parameters in neural network models: learning rate, dropout, and a number of nodes (filters for CNN and units for GRU). The effect of tuning parameters to the

models is shown in Figure 55 to 57 (Appendix B to D). Similar to XGB, learning rate determines modification of model based on updated data. Generally, the value of the learning rate depends on a number of data. If it is large, the learning rate should be large too. In contrast, the learning rate should be a small number if it is small. In this study, a number of data are small so the learning rate should be a small number in the range of 0.00001 to 0.001. Dropout is a regularization parameter to prevent overfitting by randomly eliminating nodes in each iteration. Thus, the value of dropout would consist with a number of nodes. This study suggests that the value of dropout should not be more than 0.25. The third parameter is a number of nodes for ANN, filters for CNN, and units for GRU. This parameter is varied by dropout and an amount of data. If the number of data is less than 50,000 data points, 16 to 64 is a suitable range in this study. As CNN and GRU memorize above and beyond data for classification, the number of data in memory has the effect to their performance. This study did not tune that parameter and used 15 data points above and beyond in this study.

5.3 Confusion matrix

5.3.1 Single-well test

Overall, the models predict basalt, tuff, and rhyolite accurately. Other classes such as vitrophyre and sedimentary rocks tend to be misclassified possibly due to the lack of data points and distinctive features. For example, in WO2 well in single-well test, every model predicts basalt class accurately, with over 95% of accuracy but more than 35% of conglomerate class is predicted to basalt (Figure 42). The confusion matrices of Mountain Home and Kimma in single-well test are presented in Appendix E and F.

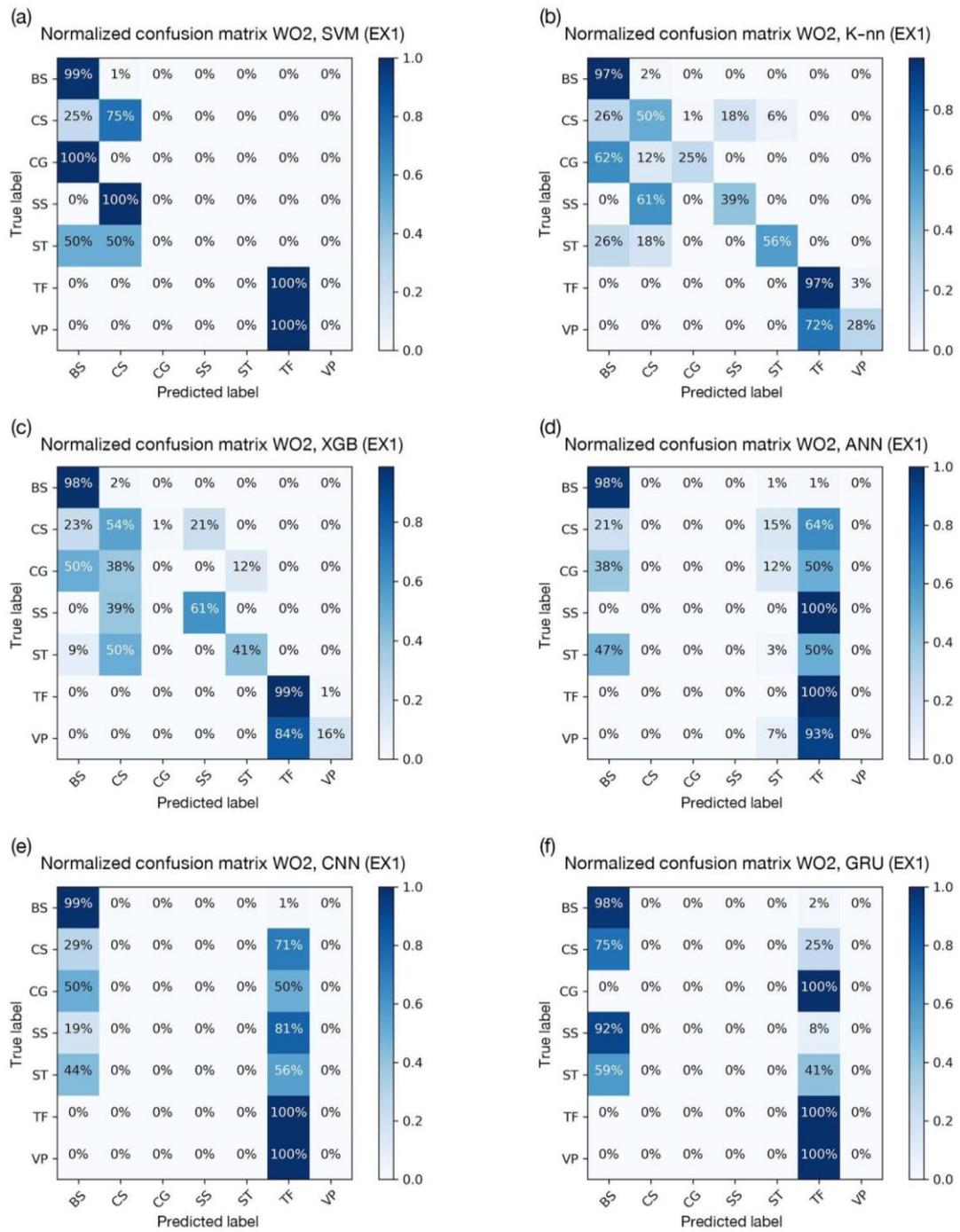


Figure 42 Normalized confusion matrix of six models in prediction of WO2 well in the single-well test. BS: basalt, CS: claystone, CG: conglomerate, SS: siltstone, ST: sandstone, TF: tuff, VP: vitrophyre.

The single-well test, training and test data are randomly selected from the same well. There are four wells: WO2, Mountain Home, Kimma, and Kimberly. In WO2 well, there are seven classes: basalt, claystone, conglomerate, sandstone, siltstone, tuff, and vitrophyre. SVM shows the high performance for classifying basalt, tuff and claystone whereas other classes such as vitrophyre, siltstone, and conglomerate are misclassified to those classes. Likewise, significantly more basalt and tuff classes than others are correctly classified by K-nn. There is moderate prediction accuracy in claystone and siltstone about 50%. Conglomerate, sandstone, and vitrophyre have low prediction accuracy. For XGB, there is almost the same pattern with K-nn but conglomerate, siltstone, and vitrophyre have lower prediction accuracy. However, XGB gives higher accuracy score in other classes than K-nn. Neural network models exhibit the lowest accuracy score for this task because there are two classes classified properly, basalt and tuff.

In Mountain Home and Kimma wells, classes are separated into two classes: basalt and sedimentary rock. The overwhelming majority of classes (more than 90%) in this wells are basalt since the models have made an attempt to give high prediction accuracy for basalt. SVM do not have capacity to distinguish the difference between these two classes with sedimentary rock being misclassified to basalt. In comparison, K-nn gives slightly accurate prediction about sedimentary rock while basalt has high precision. There is significant prediction accuracy in both basalt and sedimentary rock prediction for XGB. This would lead XGB to the best algorithm for well logging classification. Similarly to SVM, neural network models cannot recognize the difference between basalt and sedimentary rock easily, followed by giving low accuracy in prediction of sedimentary rock. However, GRU experiences the highest classification score in sedimentary rock classification among neural network models whereas it gives the lowest score in prediction of basalt.

In Kimberly well, there are three classes namely, basalt, sedimentary rock, and rhyolite. Figure 43 shows confusion matrix of each models in Kimberly well. The vast majority of classes (about 83%) are rhyolite. SVM is the best algorithm for this task because basalt and rhyolite have the highest accuracy of prediction by SVM. However, sedimentary rock still has low precision. There is high efficiency in prediction of basalt while rhyolite classification score is low in K-nn. Rhyolite is often misclassified to sedimentary rock by K-nn and sedimentary rock has low accuracy. XGB also provides high prediction accuracy in basalt and rhyolite but it gives low precision for sedimentary rock too. To compare, sedimentary rock is tended to be misclassified to basalt by K-nn and XGB whereas it is misclassified to rhyolite by SVM. The data mostly are misclassified to rhyolite being majority classes by ANN and CNN. This encourages rhyolite to have high prediction accuracy by ANN and CNN while the others have low accuracy. In comparison, GRU shows the high accuracy of prediction in basalt and rhyolite but it gives low accuracy in sedimentary rock.

5.3.2 Multiple-well test

The multiple-well test, results are almost the same as single-well test but sedimentary rock has better accuracy than previous experiments in SVM, K-nn, and XGB. This is due to the fact that this dataset has more instance of sedimentary rock. As characteristic of sedimentary rock is varied by its composition, to have more sample has encouraged the model to recognize the difference between sedimentary rock and other better than previous experiments. However, neural network models still have inefficiency in prediction sedimentary rock. The confusion matrices of every model in multiple-well test are shown in Appendix G to J.

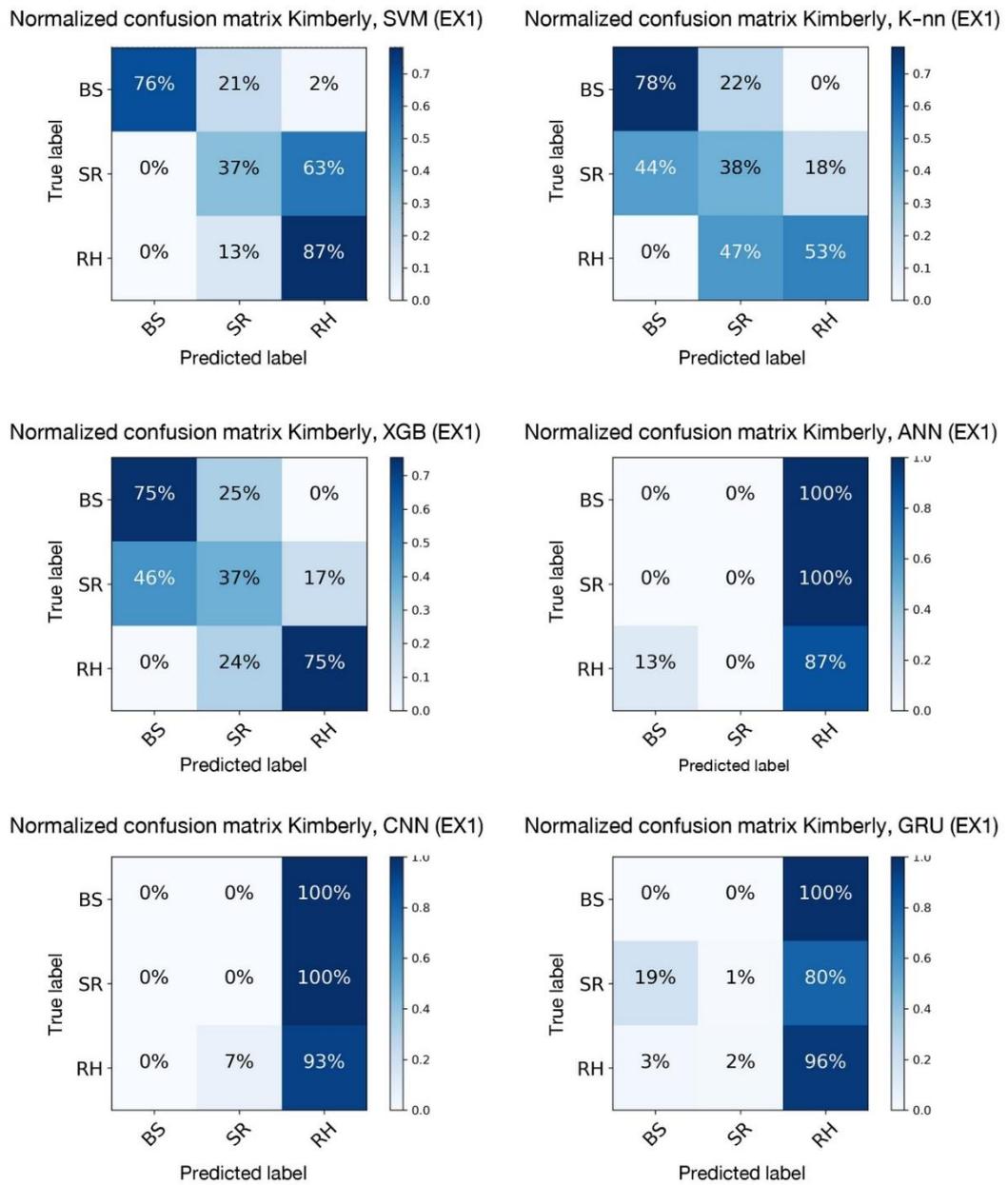


Figure 43 Normalized confusion matrix of six models in prediction of Kimberly well in the single-well test. BS: basalt, SR: sedimentary rock, RH: rhyolite.

Chapter 6 Discussion

6.1 Model performance

In this study, two experiments, namely single-well test and multiple-well test, are designed to give comprehensive results for well logging classification by deep and machine learning algorithms. In the single-well test, XGB and SVM perform best with 91% of accuracy whereas CNN exhibits worst performance at 84% (Table 4). Similarly, XGB outperforms other models with highest accuracy (87%) in the multiple-well test and CNN remains as the worst-performing model with lowest accuracy (77%). A number of previous studies e.g. Dubois et al. (2007); Konaté et al. (2015); Xie et al. (2018) apply machine learning algorithms to classify well logging data similarly to single- and multiple-well tests in this study. The previous studies use F1 score which measures accuracy while considering error but this study measures the models with accuracy. This is because every class is important equally in this study. Hence, accuracy is adequate to evaluate the performance of models. Moreover, confusion matrix which represents error and accuracy of each class is applied in this study to show the performance of the models. In this chapter, the results of this study are compared with previous studies. Even though the results cannot compare directly between accuracy and F1 score, this comparison can present the similarities, differences, and trends of each model between this study and other studies.

6.1.1 Single-well test

A comparison of results between this study and other studies in the single-well test is shown in Figure 44. The average accuracy of each model in this study is higher than 80% but the accurate performance (F1 score) in the other studies is about 60% to 80%. This is due to a difference in the number of data and a variety of classes in each study. For example, the prediction performance based on F1 score of ANN and K-nn models of sedimentary rocks in Panoma gas field, Southwest Kansas are 68% and 67%, respectively (Dubois et al., 2007). However, ANN and K-nn used in this study gives as much as 88% of classification accuracy. This is because the number of data in this study, which is about 8,000 - 20,000 data points, is 2-6 times

larger than those in Dubois et al. (2007). Moreover, well logging data in Dubois et al. (2007) are divided into eight classes, which comprise of carbonate and clastic sedimentary rocks that have similar physical and chemical properties. For example, mudstone, wackestone, and packstone are limestone classified by fabric of sediment and fossil in the rocks. Hence, composition and well logging responses of these rocks are slightly different.

Another study by Konaté et al. (2015) applies SVM, K-nn, and ANN to classify well logging data from the Dabie-Sulu terrane in China. ANN gives the highest performance in this study at 87%. As the number of data in Konaté et al. (2015) is more than 38,000 data points, deep learning algorithm such as ANN, which requires tremendous data in training step, performs well in his study. Furthermore, the rocks in the Dabie-Sulu terrane are metamorphic and igneous rocks. Well logging responses of these rocks are more unique and can be easily separable compared to those of sedimentary rocks in Dubois et al. (2007).

In addition, SVM, ANN, and XGB models are employed to classify well logging data from Ordos Basin in China by Xie et al. (2018). Similarly to this study, XGB provides the highest accurate performance at 81%. As Xie et al. (2018) has only about 1,200 data points, ANN provides lowest F1 score at 53%. The results from Xie et al. (2018) suggest that coal and carbonate rocks can be easily classified from clastic sedimentary rocks due to their composition. Clastic sedimentary rocks are mainly composed of quartz, feldspar, and clay mineral but coal and carbonate rocks are mainly composed of organic matter and calcite, respectively. As a result, well logging responses of coal and carbonate rocks are clearly different from clastic sedimentary rocks. Conversely, the classification of pebbly sandstone, coarse sandstone, medium sandstone, and fine sandstone are much harder. This is because sedimentary rocks are classified by grain size but well logging do not measure grain size directly. Not only did grain size affect well logging data, but composition also influenced well logging data. The classification of sedimentary rocks is thus more challenging than igneous and metamorphic rocks. The overall trend is about 60 - 70% of accuracy.

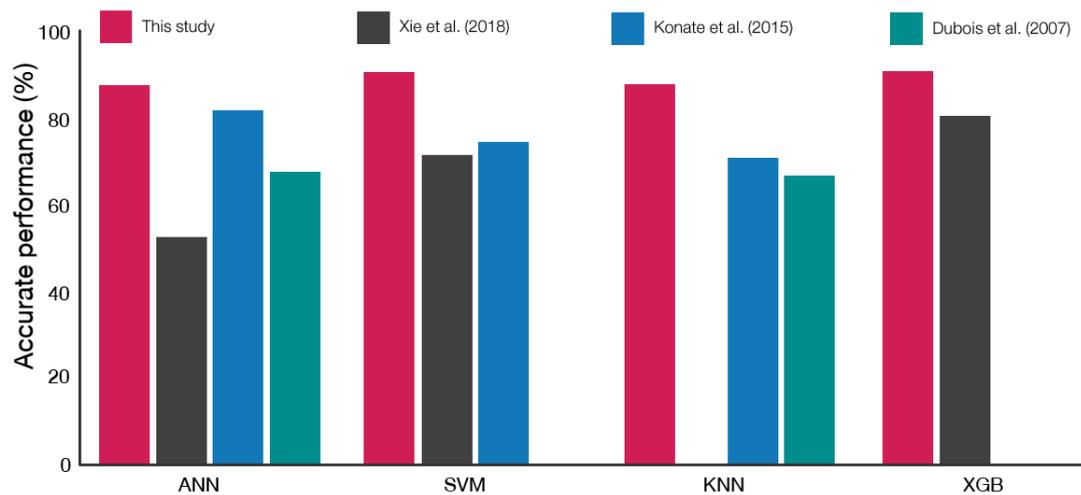


Figure 44 Comparison of accurate performance between this study and other studies in single-well test. Red color represents this study. Black refers to a study by Xie et al. (2018). Green and blue show results from Dubois et al. (2007) and Konaté et al. (2015), respectively.

6.1.2 Multiple-well test

A study by Hall and Hall (2017) uses SVM, ANN, and XGB models to classify sedimentary rocks in Panoma gas field in multiple-well test setting. The comparison of results between this study and Hall and Hall (2017) present in Figure 45. XGB gives the highest accurate performance at 63% in Hall and Hall (2017) and 87% of accuracy in this study. It is worth noting that Hall and Hall (2017) uses the same dataset as Dubois et al. (2007) but the models are employed under different data sampling. Hall and Hall (2017) combines data from eight wells as training data while using the last well as test data. In comparison, Dubois et al. (2007) select randomly 70% of whole data from every well as training data and 30% as test data. ANN in Hall and Hall (2017) provides about 56% of accurate performance while ANN in Dubois et al. (2007) gives 68% of accurate performance. Despite using the same model and dataset, results from this study and previous studies suggest that the prediction results can be significantly different due to the experimental design.

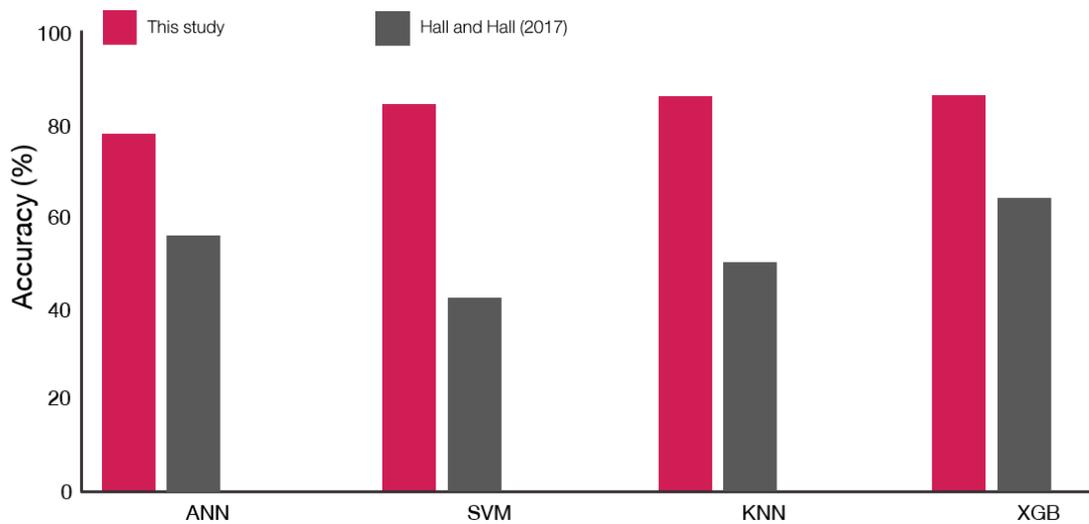


Figure 45 Comparison of accurate performance in multiple-well test between this study and Hall and Hall (2017).

In addition, accuracy results from every model in the multiple-well test are lower than those of the single-well test in this study. Figure 46 presents the comparison of accuracy between the single- and multiple-well tests. The multiple-well test is more sophisticated than single-well test because the data from different wells have more variety and complexity than data from the same well. For example, sedimentary rock in Mountain Home consists of sandstone, claystone, and breccias but sedimentary rock in Kimberly is mudstone, sandstone, and ash.

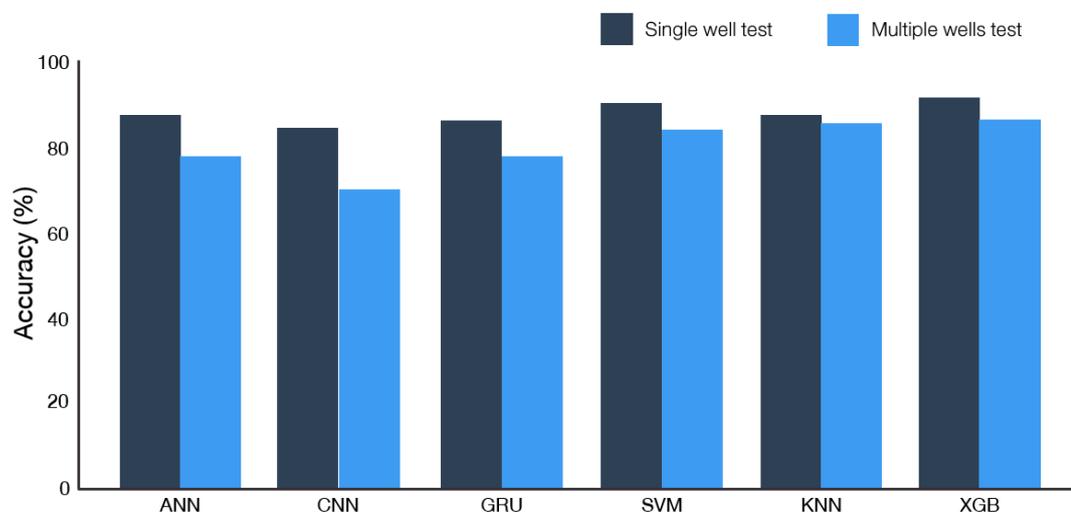


Figure 46 Comparison average accuracy of each model in this study between single-well test and multiple-well test.

Combining data from other wells can introduce noises. Foreexample, basalt and rhyolite are the different volcanic igneous rock, but their logging responses are often the same or in overlapping ranges. When basalt from other wells is combined as training data in multiple-well test, basalt basically confuses the algorithms and leads to the misclassification of rhyolite into basalt and vice versa. As a result, the accuracy of basalt in multiple-well test is increased whereas the accuracy of rhyolite is decreased (Figure 47). Moreover, the data in this study were collected by different organizations. The features or logging tools in each well vary significantly. For example, neutron log is gathered in WO2 well. The models thus have to deal with missing features when four wells are combined in multiple-well test. As a result, missing features turn sometimes into noise in the multiple-well test.

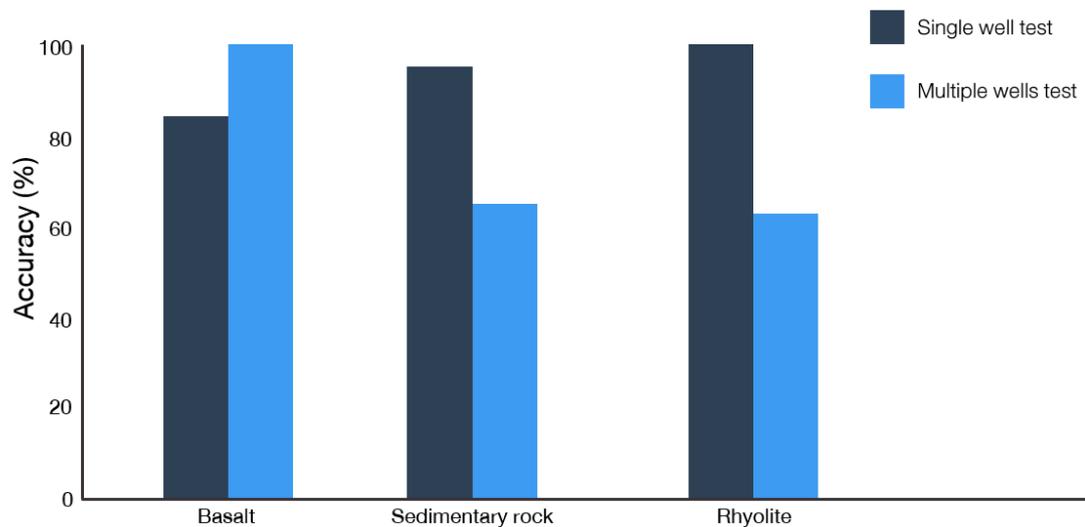


Figure 47 Comparison accuracy of each rock type in Kimberly well between single well and multiple-well test in this study. The accuracy is provided by XGB.

By merging data from different wells, two issues have emerged from the data: different sets of classes and features. Regarding the study area, lithological classification and the collected feature in each well are varied by the objectives of the study (Table 1 and 2). The models cannot predict the classes which are not included in training data and they cannot be trained with the dataset which has different sets of features. Therefore, the data should be turned into the same format in term of classes and features. The former is solved by class grouping, reduces the variety and specification of rock types. For instance, seven classes (BS, CS, CG, SS, ST, TF, and VP) in WO2 well are grouped into five classes (basalt, sedimentary rock, tuff, and vitrophyre) in multiple-well test. The latter is tackled by replacing missing value with -999.25. This value is default for absent value in Shervais (2014a, 2014b, 2014c). This has an adverse effect on the efficiency of most algorithms except XGB. For example, SVM and deep learning algorithms generate decision boundaries from training data since replacing missing value with -999.25 shifts the decision boundaries from the optimal. Likewise, it increases the distance between the test data points and the optimal training data points in K-nn. However, replacing missing value with -999.25 do not affect XGB because it can select the used features on its own. Figure 48 presents the importance of each feature which is used by XGB. The information shows that some features which is 0% of feature scores are not used for the classification. There are other techniques to fill the missing values. For example, Lopes and Jorge (2017) applies machine learning algorithms to predict the missing values from the remaining well logging data but they used 600,000 data points for regression.

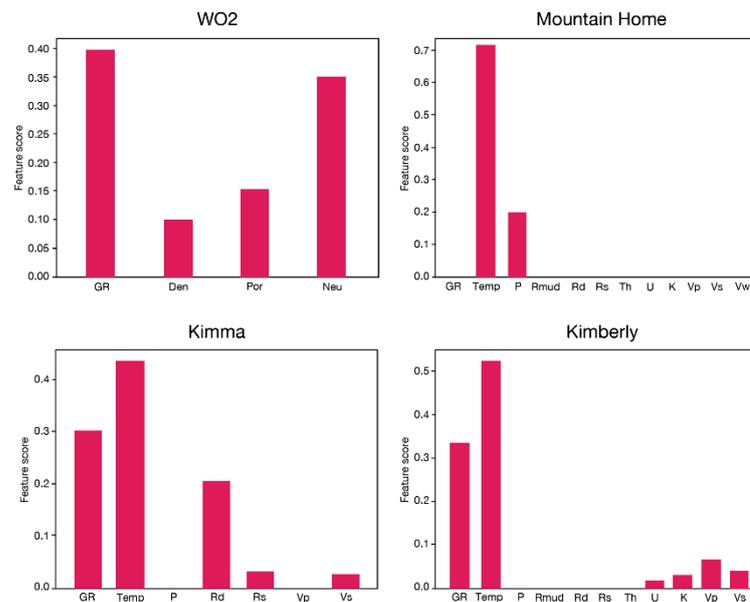


Figure 48 The used features by XGB in the single-well test. Feature score represents the importance of each feature in classification.

6.2 The effects of tuning parameters and amount of data

In this study, the models are optimized on validation data and tested on test data since the tuning parameters are tuned on validation data. However, the optimized tuning parameters on validation data are not always consistent with test data. Figure 49 shows the effect of tuning parameter on the accuracy of XGB in single-well test. Feature engineering is the other issues to improve the performance of the classifiers. For example, Bestagini et al. (2017) adopts feature augmentation. As a result, the accuracy of the classifier improves 55% to 61%.

In addition, this study attempts to normalize the data by min-max normalization. The normalized data give worse classification results than the original data. This is because the order of magnitude of each feature or logging tool significantly affects the model performance. Min-max normalization reduces the difference of feature value in each class. This allows the feature value to be more closely distributed and affects the classification boundary. The performance of K-nn

thus becomes lower because it likely uses the wrong neighboring data points to classify the data. Similarly, the accuracy of SVM and deep learning algorithms decreases because the feature values in each class is too close to generate optimal decision boundaries. However, the performance of XGB remains high because the ensemble decision tree uses the feature values sequentially. In contrast to other models, XGB does not use every feature at the same time.

The most important parameter affecting the performance of each model in this dataset is a number of data. As a number of igneous rocks are significantly more than sedimentary rocks as called class imbalance, the models give the accuracy of igneous rocks higher than sedimentary rocks. Moreover, the effect of imbalance classes causes the classifiers to recall the major classes more so than the minor classes. As result, the minor classes are misclassified into the major classes. This effect influences the classifiers using the decision surfaces for classification such as SVM and deep learning algorithms than the other. This is because decision surfaces which are generated from small data points are not fully comprehensive the distribution of classes.

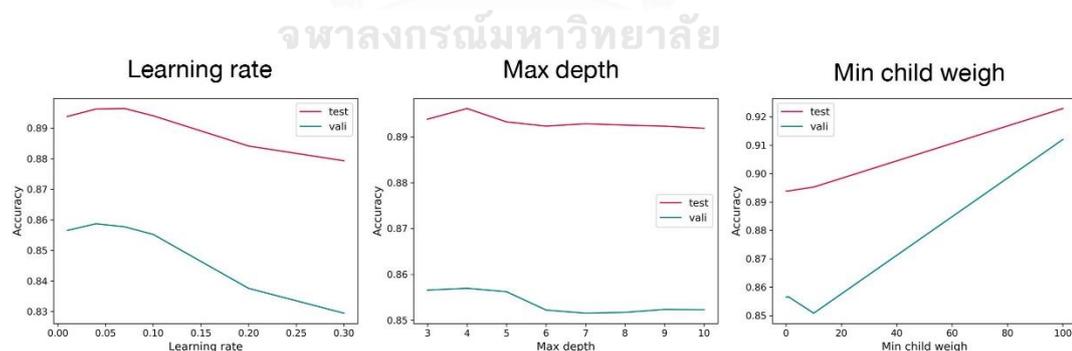


Figure 49 The effect of tuning parameter to the accuracy of XGB in single-well test. Red lines represent test subset and green lines represent validation subset.

Each of classifiers has different benefits and drawbacks and they are suitable with various dataset. K-nn is suitable for the data which are no explicit knowledge and there is no time spent in the training phase for K-nn. K-nn is over-reliant on training data because it does not assign any assumption into the data (Glowacz and Glowacz, 2016). There is no feature weigh function and noise filter for K-nn. Hence, feature selection and cleaning data greatly influence the performance of K-nn than other algorithms. As K-nn employs feature similarity to classify data, classes which have nearby feature such as basalt and rhyolite or claystone and other sedimentary rocks could confound it. For further study, feature weighting should be applied for K-nn. K-nn considers every feature equally but, in real task, some features are not relevant to the classification. Consequently, eliminating irrelevant and noisy features could improve the classification accuracy of K-nn. For example, a study by Panday et al. (2018) demonstrates that K-nn with feature weighting algorithm gives better accuracy score than original K-nn, especially a dataset having many features.

SVM is a generalized algorithm because it considers merely support vectors (Smirnof et al., 2008). By doing this, SVM can avoid noise and ambiguous data as well as soft margin encourages the model to more generalize. For example, there are two subgroups of rhyolite in Kimberly well, upper rhyolite and lower rhyolite, which have slightly different composition. XGB and K-nn think whether they are different classes and misclassify upper rhyolite to basalt but SVM predicts them as rhyolite. However, SVM often misclassifies the minor classes, which have a small number of data. SVM classify data by drawing decision boundaries (Meyer et al., 2003) since the decision boundaries generated from small data are harder to be the accurate decision boundaries than the decision boundaries generated from big data. Moreover, SVM with soft margin allows some data points to be misclassified. This reduces a number of data in the minor classes considered to draw decision boundaries of the

minor classes. As a result, the large number of minor classes are predicted to the major classes by SVM.

In this study, three deep learning algorithms are applied on this dataset: ANN, CNN, and GRU. ANN gives the highest prediction accuracy among three deep learning algorithms whereas CNN exhibits the lowest. Due to class imbalance in training data, deep learning algorithms do not have the capacity to classify the minor classes properly. Figure 50 shows a comparison of accuracy between ANN and XGB with the size representing a number of data. ANN gives 0% of accuracy in classes which have a number of data lower than 500 data points. In comparison, basalt and tuff being the vast majority of data have high prediction accuracy and deep learning algorithms tend to predict other classes to the majority (Figure 42). Neural network models attempt to maintain the prediction accuracy as high as they can while saving the computational cost. They thus recall the major classes dramatically to make high accuracy when data are significantly imbalanced. As CNN gives worse efficiency than ANN, this indicates that feature extraction is not necessary for well logging data. Each logging tool is designed to detect characteristic of subsurface rock since feature weight is adequate for well log interpretation. Furthermore, GRU and ANN show similar results because there is less relationship in time series. SRP is volcanic field which igneous rock could inject to anywhere based on weak zone and tectonic processes so there is less gradually change which GRU can show the high potential.

Deep learning algorithms are easier to overfit with the major classes than the other algorithms. This is because nodes are adjusted to be fit with data in each iteration. The major classes can be passed into the node than the minor classes. As a result, every node is adjusted for major classes. To prevent overfitting, dropout, which randomly omits some nodes in each iteration of training step, is applied in this study (Srivastava et al., 2014). Although deep learning algorithms perform better, they still misclassify the minor classes. For CNN and GRU, another parameter which affects

efficiency of model is a number of memorized data but it is not tuned in this study. For further study, down-sampling should be tried to reduce class imbalance. This might improve the performance of deep learning algorithms. Moreover, hybrid models for well logging classification should be used. The hybrid model is the new machine learning algorithm which combines the advantages of more than one models since its performance might be better than the ordinary model. For example, Zhu et al. (2016) and Zhu et al. (2018) apply the hybrid model between ANN and random forest to predict permeability and total organic carbon using well logging data, respectively. Results show that the hybrid model gives a lower error than both ANN and random forest.

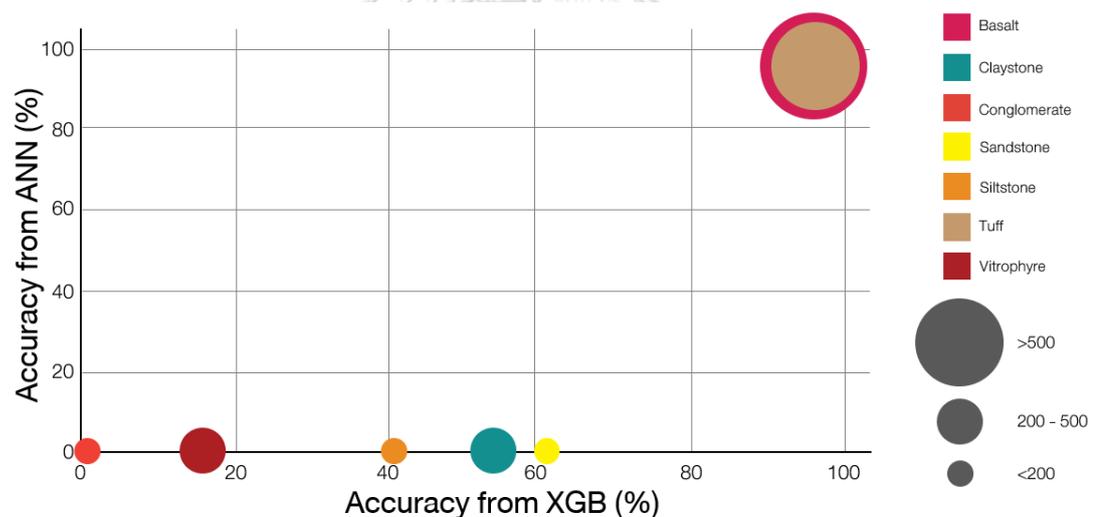


Figure 50 A comparison of accuracy between ANN and XGB with a number of data represented by size of spot. ANN cannot classify minor classes correctly while XGB can.

Results from Hall and Hall (2017); Xie et al. (2018); and this study present that XGB or Gradient tree boosting exhibits the highest the classification accuracy for lithological classification using well logging data. This is because XGB can select the used features in its own without human influence. Figure 51 shows the importance

of each feature for this dataset in multiple-well test. Gamma ray, temperature, and neutron logs are important features to classify the rocks in Snake River Plain in which is dominated by volcanic rocks. As humans do not have the capacity to know which features are appropriate for each algorithm, it is better if algorithms choose the used features by itself. As a result, XGB can recognize that -999.25 is missing value for this dataset and it does not use this value in classification. Moreover, XGB classifies the data by decision trees which generate if-clause rules from characteristics of training data to classify test data. This is close to how humans classify the data since XGB is appropriate with well logging data which is invented for the human to classify rock types.

6.3 Feature importance and ranges

Feature importance refers to the frequency of features used to classify data. Figure 52 shows the average feature importance scores from XGB. High feature importance score indicates that this feature is frequently used by XGB. Feature ranges indicates the values of the feature used to determine each class by XGB. Table 6 shows various ranges used by each well logging tool (feature) to classify rocks with XGB model in single-well test. Regarding Figure 17, basalt and tuff are two classes classified by XGB accurately, with having accuracy more than 95%. Conglomerate and vitrophyre have the lowest (<20%) classification accuracy. The accuracy of sandstone, siltstone, and claystone given by XGB are in a range of 40% to 60%. For igneous rocks, our results are compared with two works, namely Berendsen et al. (1988) and Huang et al. (2015). The former studies the well logging data in the Midcontinent Rift System, Kansas, the USA and the latter studies well logging data in the Anda Sag, the Songliao Basin.

Table 8 Range of well logging values in each rock from WO2 well in single-well test by XGB compared to the other works.

Rock types	Gamma ray (API)	Density (g/cm ³)	Neutron (API)	Porosity (%)	Study
Basalt	10.59 - 15.15	2.62 - 2.88	464.27-604.58	7.40 - 24.51	<i>This study</i>
	10 - 40	> 2.9	-	10 - 20	Berendsen et al. (1988)
	37.5 - 50	≈ 2.75	-	-	Huang et al. (2015)
Tuff	>119.17	<2.24	>837.25	≈ 0	<i>This study</i>
	≈ 150	2.4 - 2.6	-	-	Huang et al. (2015)
Vitrophyre	>119.17	2.24 - 2.62	604.58-837.25	>24.51	<i>This study</i>
Conglomerate	15.15-119.17	2.62 - 2.88	>464.27	7.40 - 24.51	<i>This study</i>
	70 - 80	-	-	-	Whittemore et al. (1997)
Sandstone	15.15-119.17	2.24 - 2.62	<464.27	>24.51	<i>This study</i>
	75 - 80	-	-	-	Whittemore et al. (1997)
	25 - 50	1.67 - 2.10	-	-	Lin et al. (2018)

Table 9 Range of well logging values in each rock from WO2 well in single-well test by XGB compared to the other works (continued).

Rock types	Gamma ray (API)	Density (g/cm ³)	Neutron (API)	Porosity (%)	Study
Sandstone	15.15-119.17	2.24 - 2.62	<464.27	>24.51	<i>This study</i>
	75 - 80	-	-	-	Whittemore et al. (1997)
	25 - 50	1.67 - 2.10	-	-	Lin et al. (2018)
Siltstone	15.15-119.17	2.24 - 2.62	604.58-837.25	>24.51	<i>This study</i>
	80 -130	-	-	-	Whittemore et al. (1997)
	100 - 150	2.56 - 2.78	-	-	Lin et al. (2018)
Claystone	15.15-119.17	2.24 - 2.62	<464.27	>24.51	<i>This study</i>
	100 -150	-	-	-	Whittemore et al. (1997)
	125 - 150	2.12 - 2.54	-	-	Lin et al. (2018)

This study compares the range of gamma ray used by XGB and by human because generally, human uses gamma ray to calculate the composition of rocks in order to determine rock types. Moreover, the gamma ray is the most frequently used by XGB to classify well logging data as shown in Figure 52. XGB uses gamma ray in ranges of 10.59 to 15.15 API to determine basalt in this study. Berendsen et al. (1988) uses gamma ray in ranges of 10 to 40 API to determine basalt while Huang et al.

(2015) uses 37.5 to 50 API. XGB uses gamma ray in the same range of Berendsen et al. (1988) but uses in the different range of Huang et al. (2015). Basalt in Huang et al. (2015) from the Anda Sag is disturbed by magma differentiation so the composition of basalt in Huang et al. (2015) is more felsic. Biotite which has potassium being the source of gamma is increased in basalt's Huang et al. (2015) by magma differentiation. In contrast, there is no magma differentiation in our area. Consequently, the range of gamma ray in Huang et al. (2015) is higher than this study.

The gamma ray of tuff used by XGB in this study is more than 119 API whereas Huang et al. (2015) uses 150 API to classify tuff. Figure 51 shows well logging marks of volcanic rocks and types of interface in volcanic rocks determined by Huang et al. (2015). Tuff in Huang et al. (2015) is determined as a volcanic bed and there is only one event producing tuff in the Anda Sag. In comparison, there are many volcanic eruptions producing tuff in the Snake River Plain since the composition of tuff in this study is more variety than Huang et al. (2015).

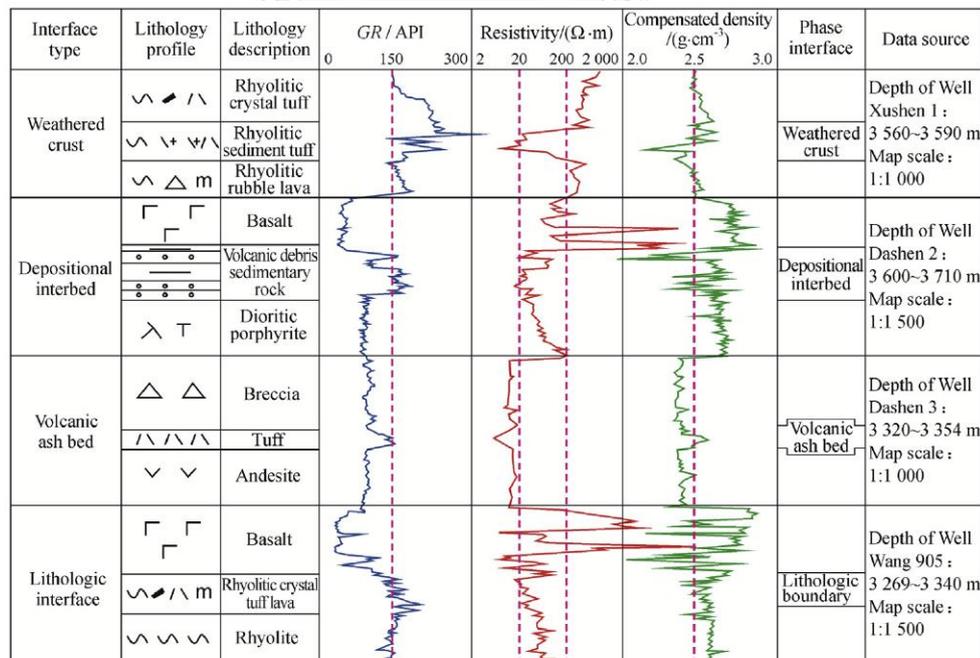


Figure 51 Well logging marks of volcanic rocks in the Anda Sag, the Songliao Basin (Huang et al., 2015).

As XGB give low accuracy in classification conglomerate, well logging values used to determine conglomerate by XGB are not reliable. The other sedimentary rock such as claystone, siltstone, and sandstone have moderate accuracy score, followed by being in ranges of 50% to 60%. In sedimentary rocks, the well logging values from Dakota aquifer (Whittemore et al., 1997) and gas hydrate reservoir in the Muli area of Qinghai, China (Lin et al., 2018) are used to compare with feature ranges from XGB. The maximum and minimum ranges of gamma ray of claystone, siltstone, and sandstone are the same as 15.15 to 119.17 API. As sedimentary rocks are ambiguous for XGB, the values of each well logging tools (features) are in the same or close ranges as shown in Table 6. However, the gamma ray of sedimentary rocks used by XGB is close to the other studies. Even though XGB cannot clearly distinguish sedimentary rocks, the used features are similarly to human classification.

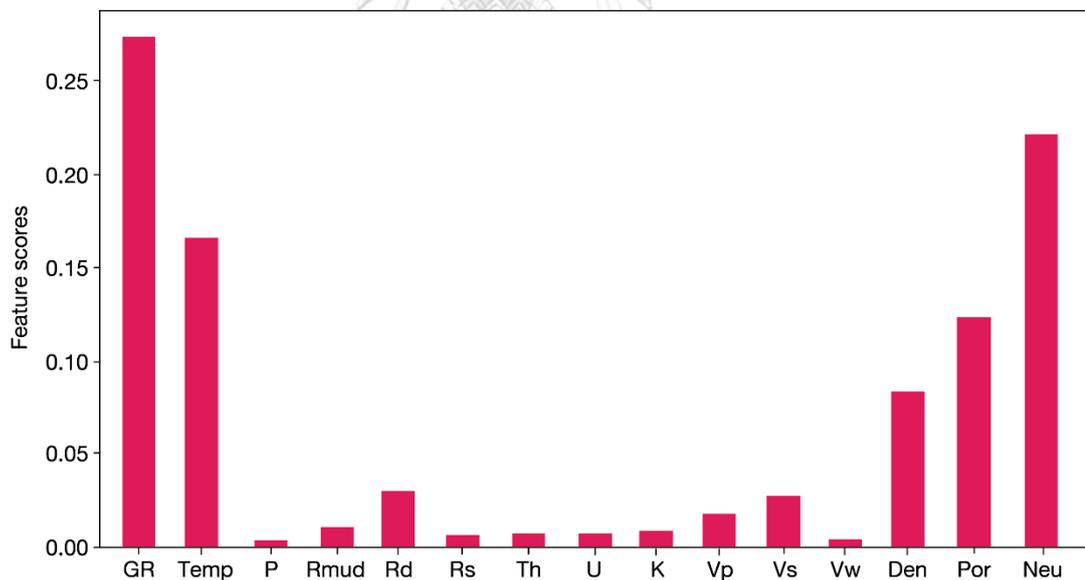


Figure 52 Average feature importance scores of each feature or type of well logs in XGB algorithm from multiple-well test.

Chapter 7 Conclusions

7.1 Conclusions

Six machine learning algorithms are employed to classify lithology of well logging data acquired from four geothermal wells in the Snake River Plain, Idaho, US. The experiments are divided into two experimental conditions, single- and multiple-well tests, in order to evaluate the models in different use cases. The model parameters are optimized by hyperparameter tuning method on validation dataset. To prevent the effect of overfitting, data are randomly divided five times into training, validation, and test set respectively. The models are further trained and tested five times with different dataset to obtain average accuracy for both experimental conditions.

Results suggest that Extreme gradient boosting (XGB), which is a decision tree-based model, shows the highest accuracy (91% and 87%) in lithological classification in both single- and multiple-well tests. This is due to the fact that XGB has feature selection in its algorithm and operates similarly to human decision. By doing this, well logging data which design for human to interpret lithology of subsurface rock has been efficiently classified lithology by XGB. In addition, XGB is the only algorithm in this study that can handle missing data during training by classifying the missing data to into the left or right node, which can minimize the loss function. However, the other classifiers excepted K-nn use feature weight function to deal with this problem instead.

Other algorithms such as Support vector machine (SVM) and K-nearest neighbour (K-nn) show satisfactory results even their accuracy are lower than XGB. The performance of deep learning algorithms such as Artificial neural network (ANN), Convolutional neural network (CNN), and Gated recurrent unit (GRU) models are inferior to XGB as they tend to require balanced data. This study suggests that deep

learning algorithms are not the most suitable algorithm for lithology classification from well-logging data.

Furthermore, the classes of this dataset are significantly imbalanced. The vast majority of data are basalt and rhyolite whereas sedimentary rock is the minority. Machine learning algorithms thus have high potential to classify basalt and rhyolite precisely but sedimentary rock has low prediction accuracy in this study. This is because sedimentary rocks have closely similar well logging response with having a small amount of data. Moreover, sedimentary rocks are also classified by grain size and composition but there are no well logging tools detecting grain size directly. Consequently, to specify sedimentary rock is another challenge of lithological classification using machine learning algorithm.

In addition, machine learning algorithms spend a second to classify well logging data in test step for about thousands of data points in both single- and multiple-well test. For training step, the models except K-nn spend about half of hour in single-well test (5,000 – 15,000 data points) and an hour in multiple-well test (about 40,000 data points). Even though machine learning algorithms cannot give a perfect result in well logging classification, they can help human reduce time-consuming in well logging classification.

7.2 Recommendations

For further study, hybrid models should be developed and feature engineering should be applied to improve the accuracy and efficiency of well logging classification. Moreover, standard practice of well logging data collection should be followed for lithological classification by machine learning algorithms. For example, gamma ray, neutron, and temperature logs are necessary data for this dataset. Conversely, this study suggests that pressure and water wave velocity data are unnecessary logging tools in rock classification. These key insights can be used to

improve the well logging data collection planning and procedure as well as reduce the budget and time for well logging surveys in the future.

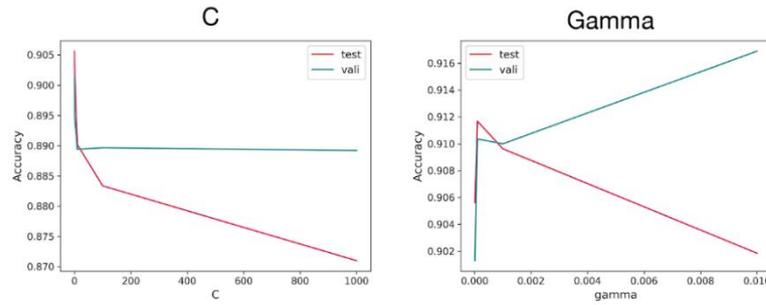
In addition, the same terminology of rock classification within the dataset should be applied. This is to reduce classification ambiguity and misclassification. Sedimentary rocks in this study are significantly misclassified due to the two factors. Firstly, the feature values of all sedimentary rocks are very similar, almost inseparable, causing incorrect decision boundaries. Secondly, the well logging data is provided with different terminologies to describe the same rock type. This issue can be improved by using the same terminology for entire dataset.

This study can also serve as a pilot study for well logging data in Thailand. Well logging data in Thailand mainly comprises of sedimentary rocks such as sandstone, shale, and siltstone. Results from this study show that machine learning algorithms give low accuracy in classifying sedimentary rock. However, increasing the number of data can improve the performance of machine learning algorithms. Consequently, the large number of data with balanced class should be input into machine learning algorithms to learn the characteristics of each rock type in Thailand. Machine learning algorithms can reduce time-consuming and improve classification efficiency in well logging interpretation. This help improve natural resource exploration and production.

APPENDIX A

The effect of tuning parameters

Single-well test



Multiple-well test

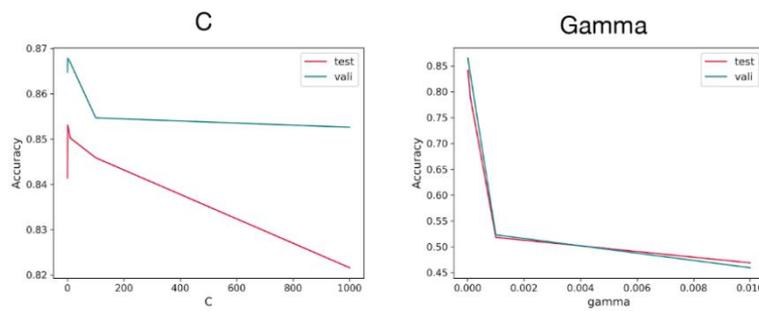
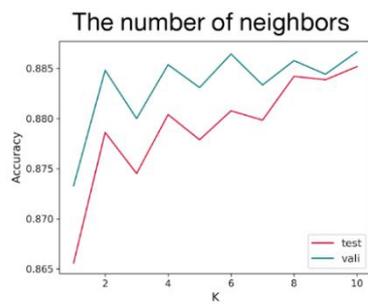


Figure 53 The effect of tuning parameters in SVM.



Single-well test



Multiple-well test

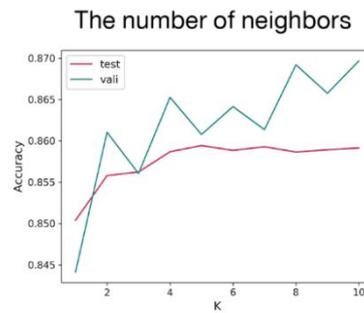


Figure 54 The effect of tuning parameter in K-nn.

APPENDIX B

The effect of tuning parameters (continued)

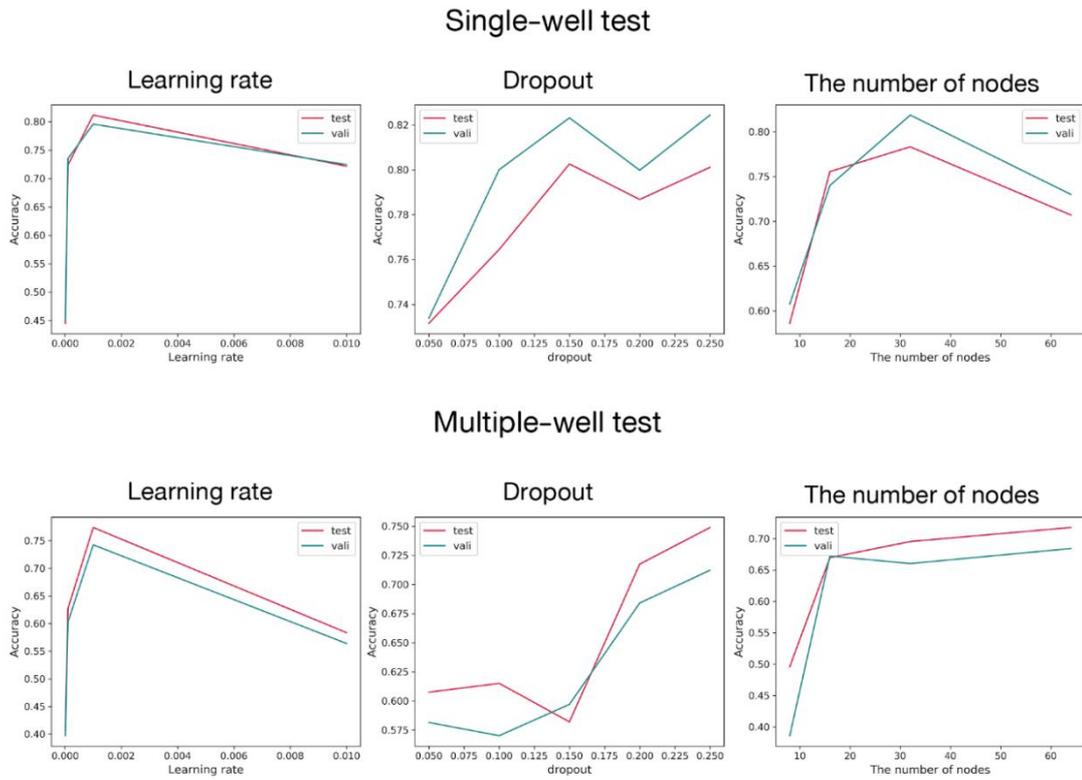


Figure 55 The effect of tuning parameters in ANN.



APPENDIX C

The effect of tuning parameters (continued)

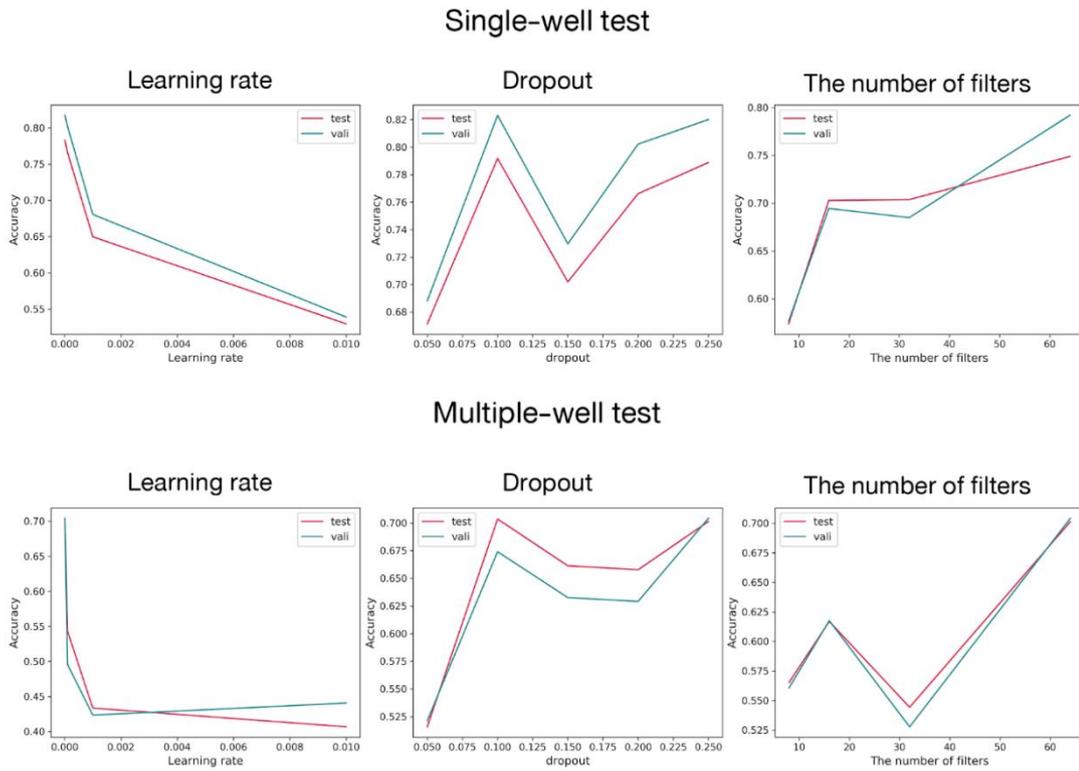


Figure 56 The effect of tuning parameters in CNN.



APPENDIX D

The effect of tuning parameters (continued)

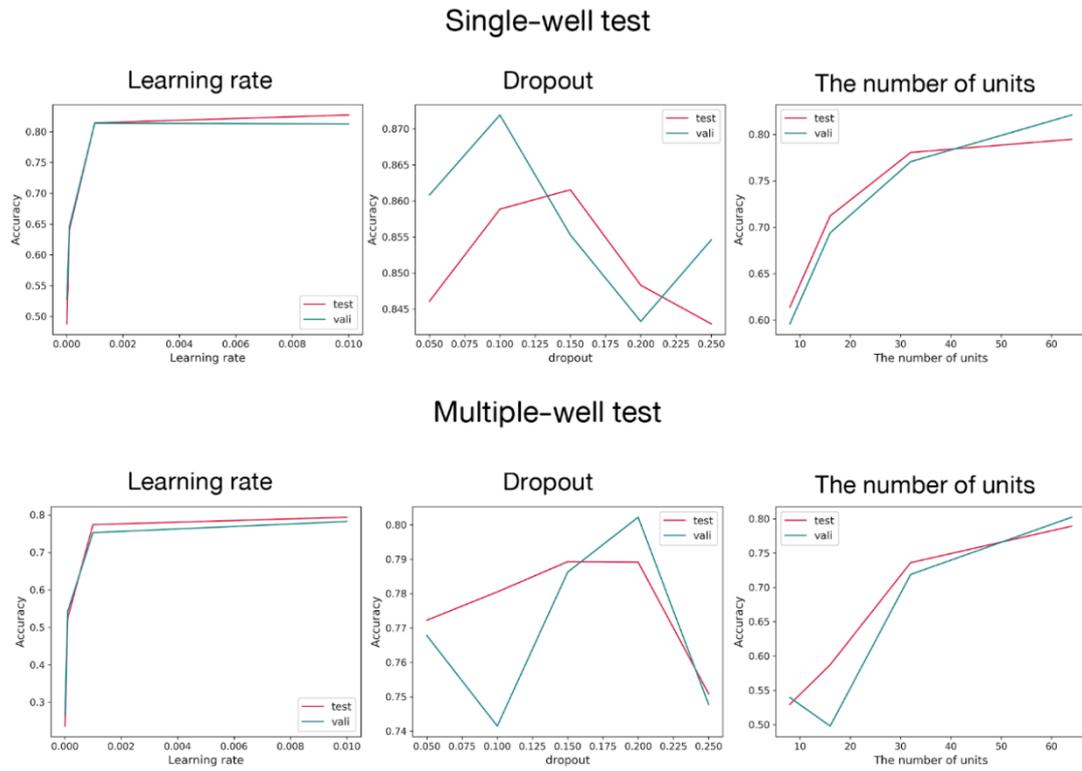


Figure 57 The effect of tuning parameters in GRU.



APPENDIX E

Confusion matrix of each model in a single-well test with Mountain Home well.

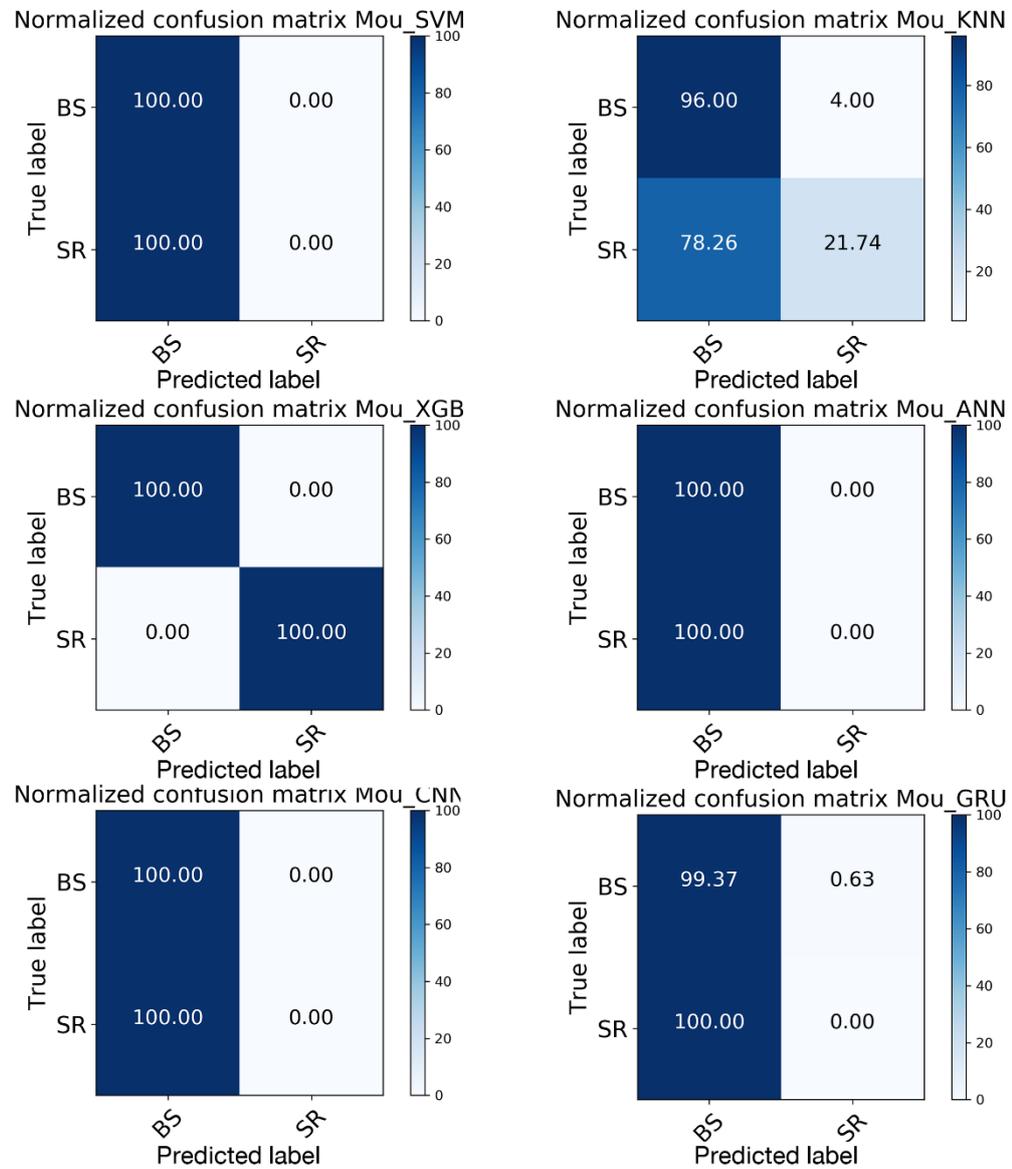


Figure 58 Normalized confusion matrix of six models in prediction of Mountain Home well in the single-well test. BS: basalt, SR: sedimentary rocks.

APPENDIX F

Confusion matrix of each model in a single-well test with Kimma well.

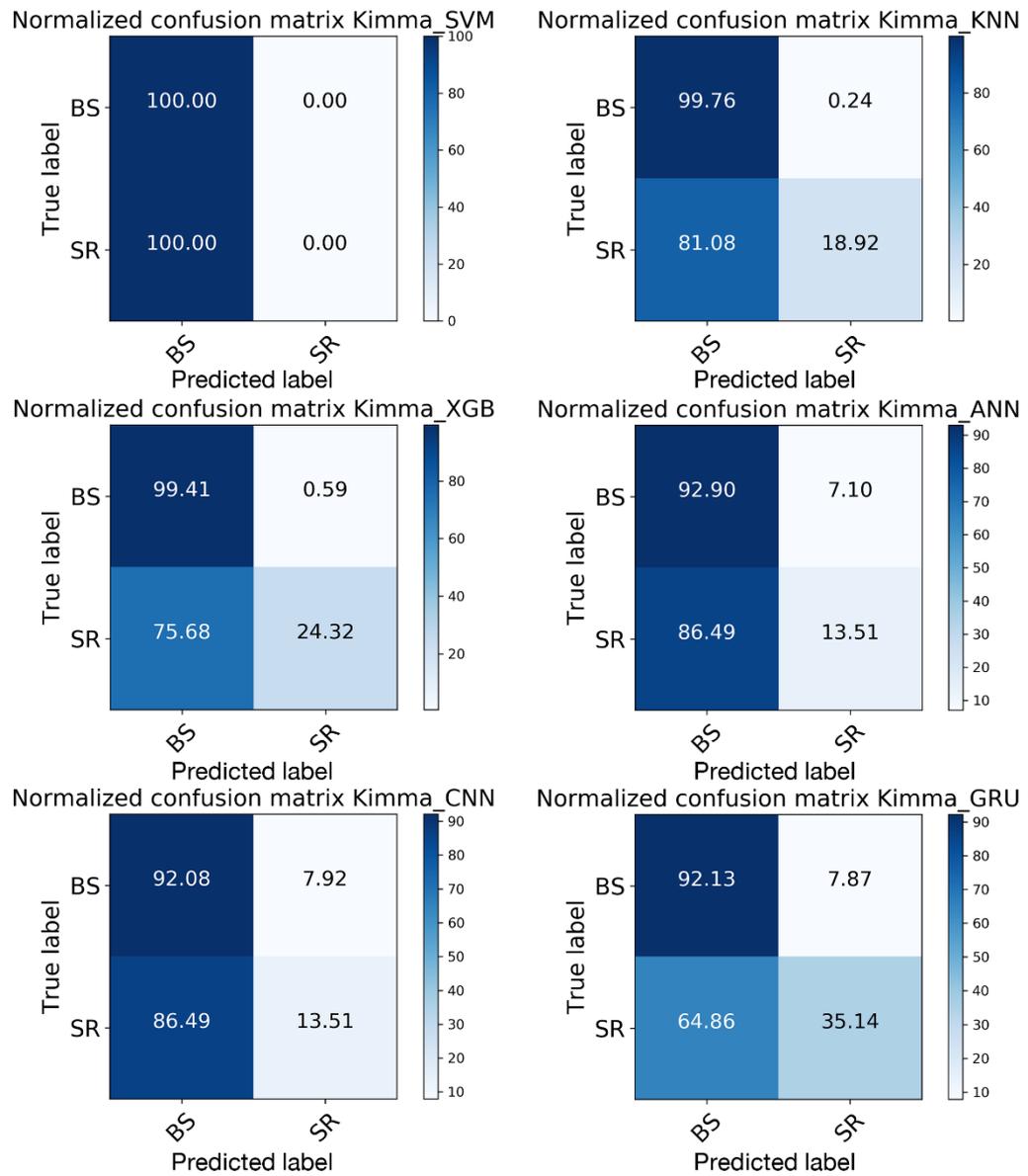


Figure 59 Normalized confusion matrix of six models in prediction of Kimma well in the single-well test. BS: basalt, SR: sedimentary rocks.

APPENDIX G

Confusion matrix of each model in a multiple-well test with WO2.

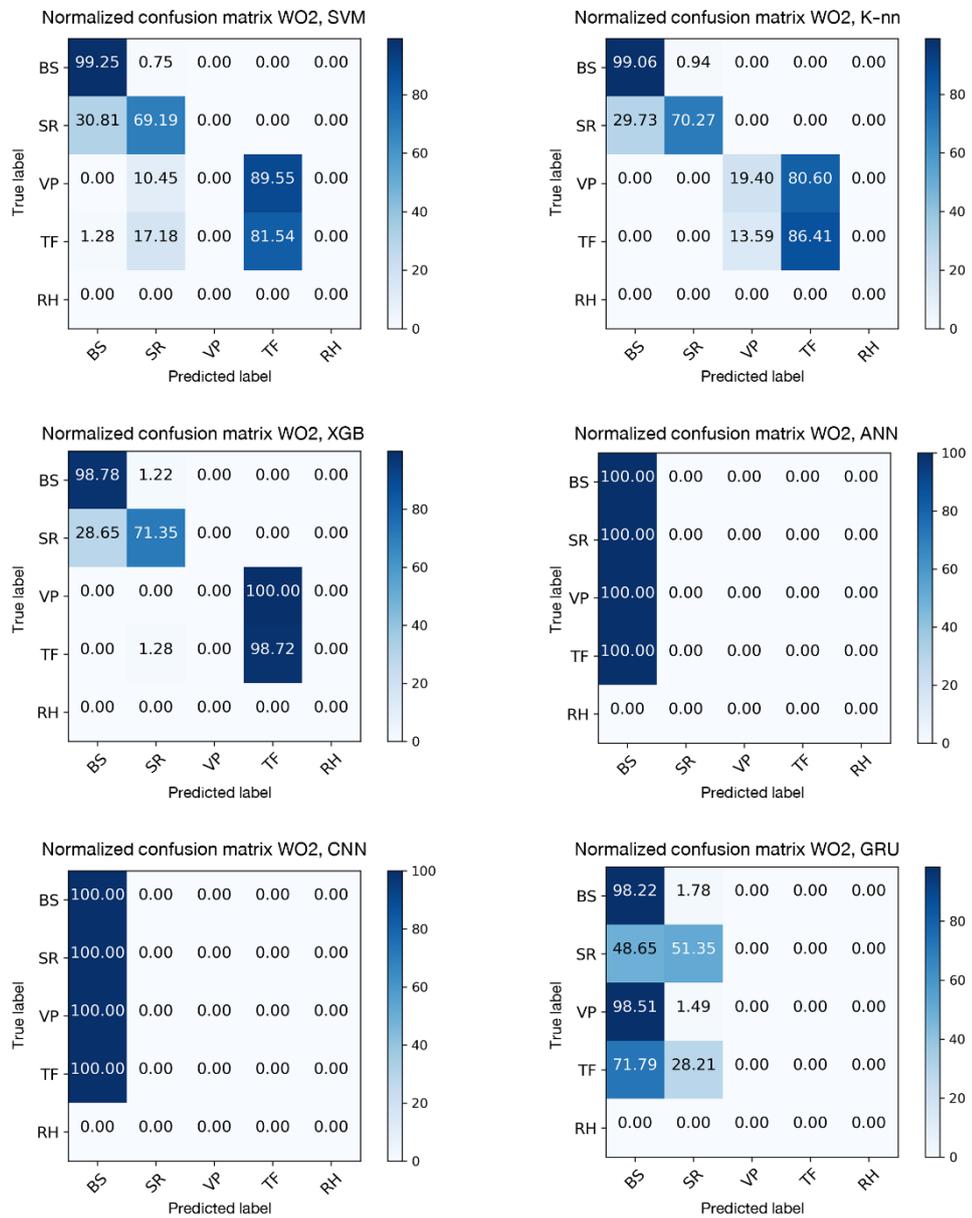


Figure 60 Normalized confusion matrix of six models in prediction of WO2 well in the multiple-well test. BS: basalt, SR: sedimentary rocks, VP: vitrophyre, TF: tuff, RH: rhyolite.

APPENDIX H

Confusion matrix of each model in a multiple-well test with Mountain Home.

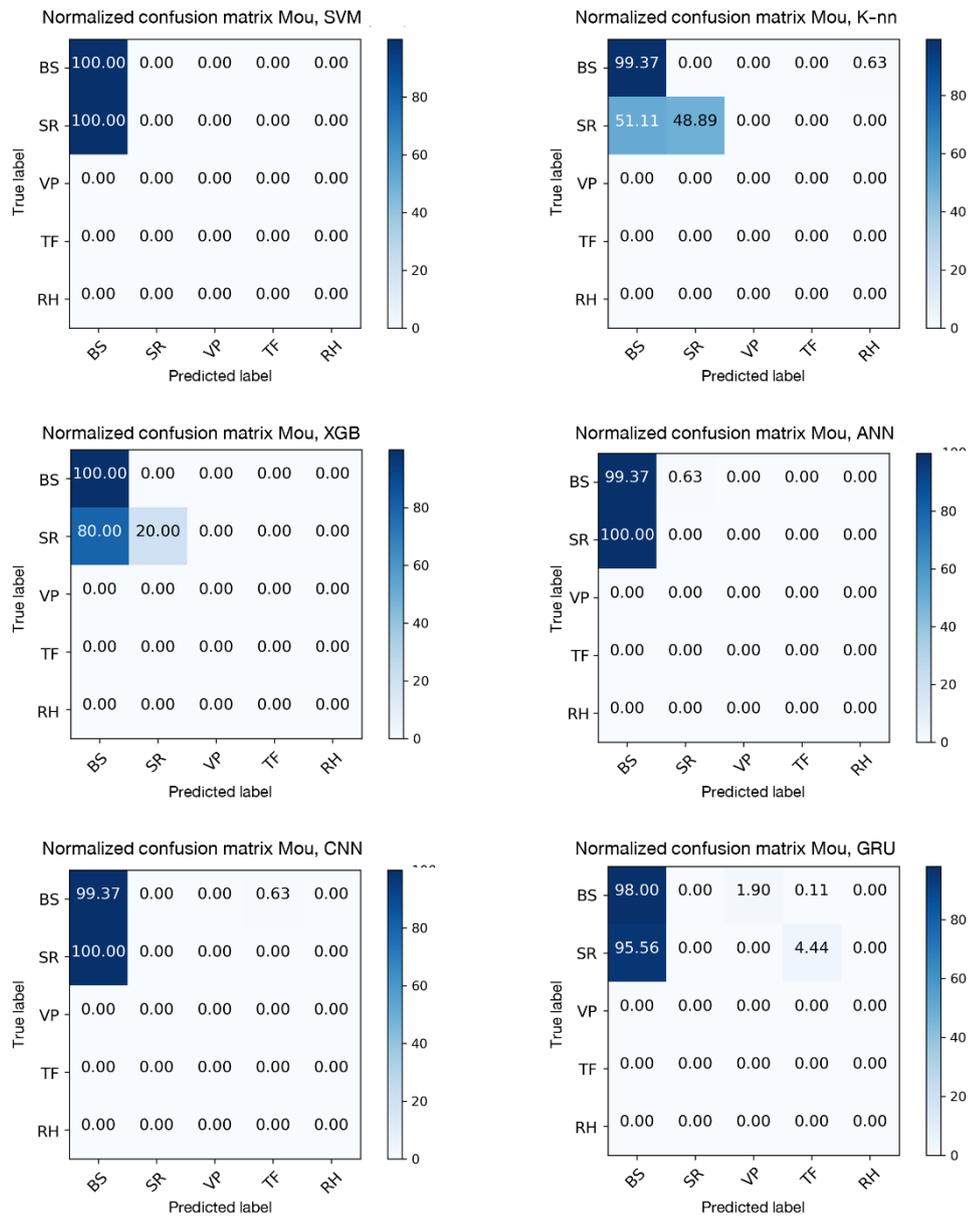


Figure 61 Normalized confusion matrix of six models in prediction of Mountain Home well in the multiple-well test. BS: basalt, SR: sedimentary rocks, VP: vitrophyre, TF: tuff, RH: rhyolite.

APPENDIX I

Confusion matrix of each model in a multiple-well test with Kimma.

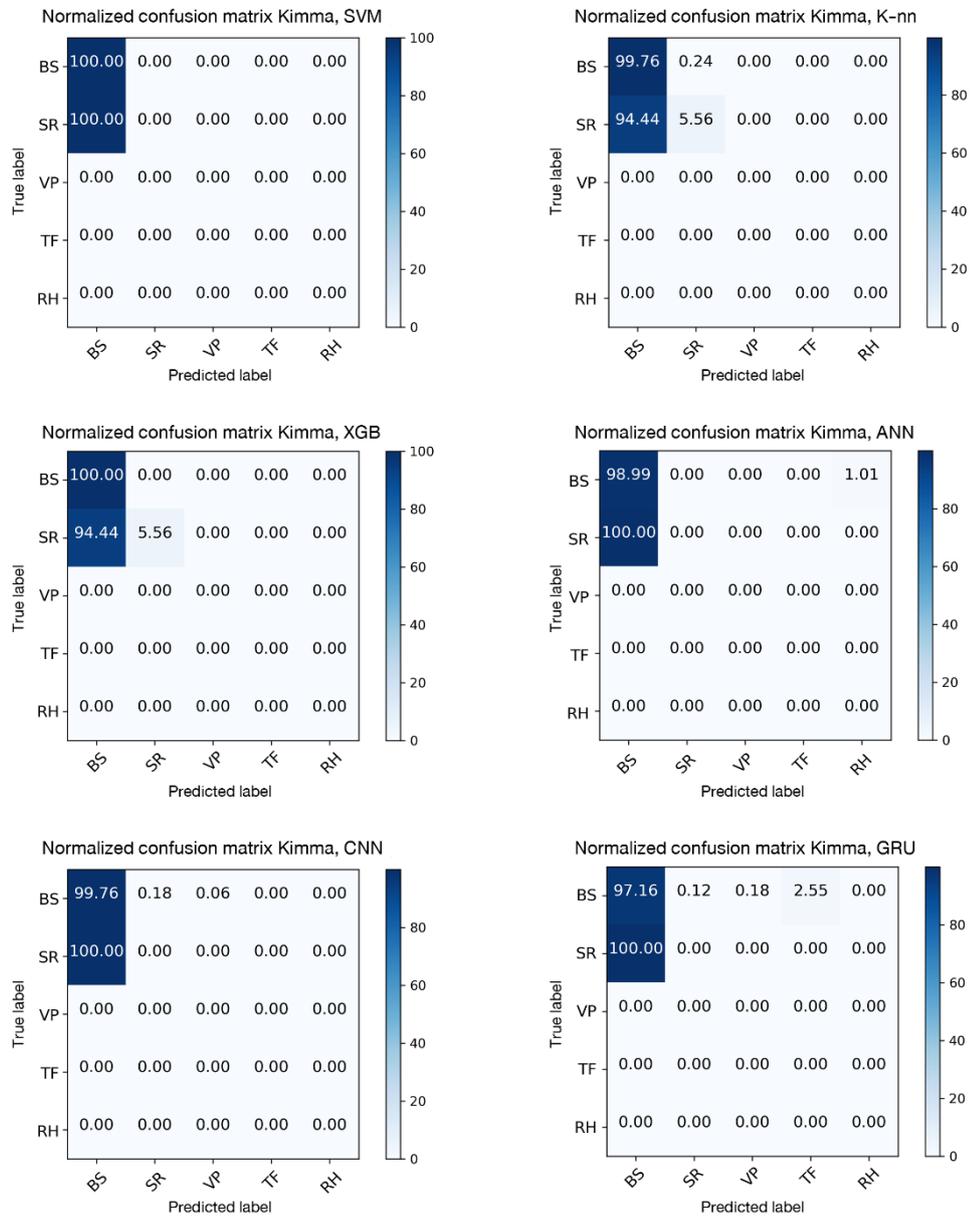


Figure 62 Normalized confusion matrix of six models in prediction of Kimma well in the multiple-well test. BS: basalt, SR: sedimentary rocks, VP: vitrophyre, TF: tuff, RH: rhyolite.

APPENDIX J

Confusion matrix of each model in a multiple-well test with Kimberly.

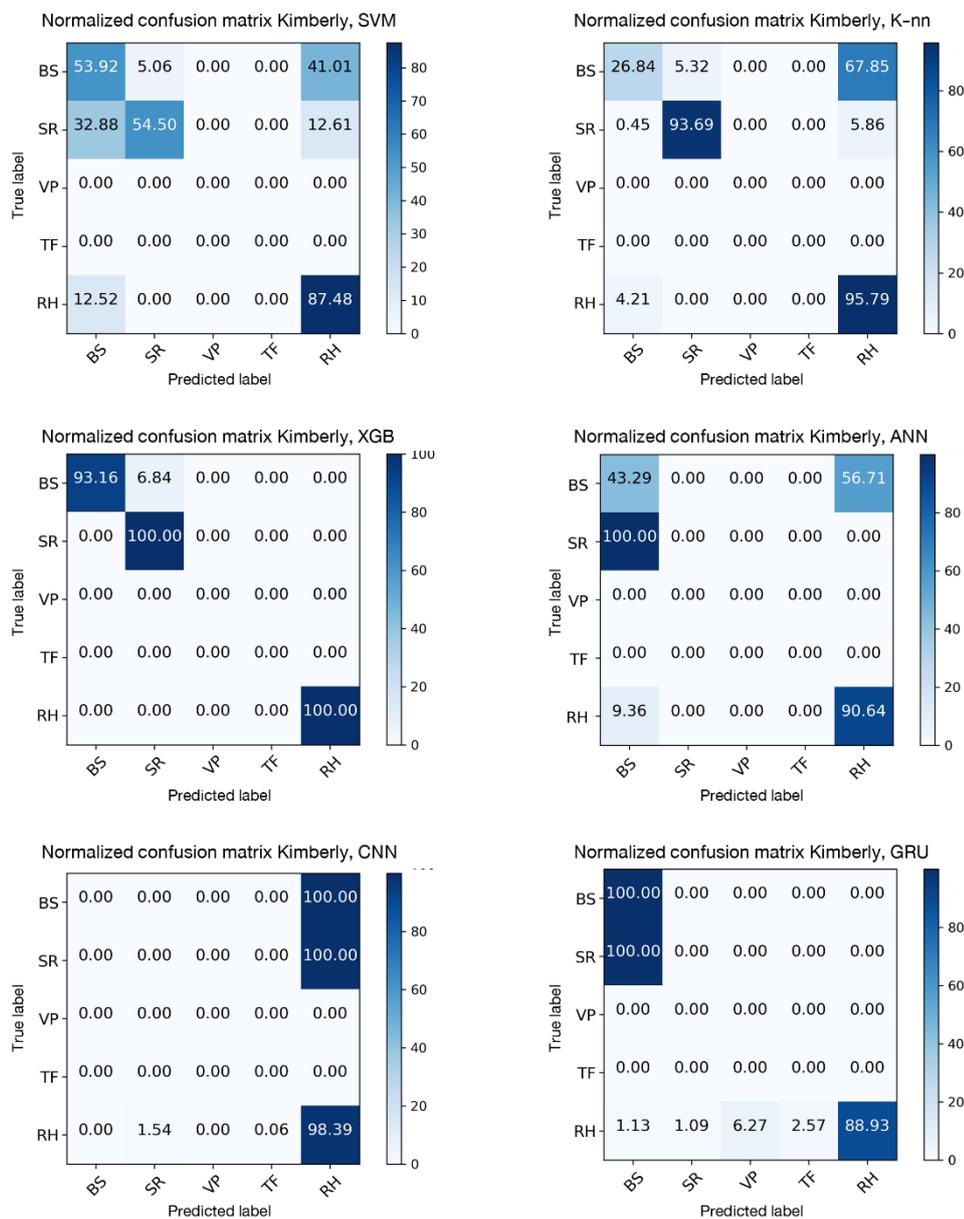


Figure 63 Normalized confusion matrix of six models in prediction of Kimberly well in the multiple-well test. BS: basalt, SR: sedimentary rocks, VP: vitrophyre, TF: tuff, RH: rhyolite.

APPENDIX K

Raw data of well logging data from the Snake River Plain used in this study are published on Github: <https://github.com/Worapop/Well-logging-data-from-Snake-River-Plain>

Depth m	GR API	TEMP1 C	TEMP2 C	P MPa	Rmud Ohmm	Rd Ohmm	Rs Ohmm	Th ppm	U ppm	K %	Vp m/s	Vs m/s	Vw m/s	INDEX Rock types
1173.5	17.0524	91.6933	91.7334	10.4903	1.77932	2.95901	3.61663	2.8906	-999.25	0.660027	2634.77	1939.17	1315.71	Basalt
1173.6	17.5466	91.7012	91.7466	10.4913	1.77938	3.04435	3.79231	3.04549	-999.25	0.73122	2634.77	1938.44	1314.36	Basalt
1173.7	17.6674	91.71	91.7466	10.4922	1.77915	3.26178	4.07538	3.20912	-999.25	0.810748	2633.43	1936.26	1314.03	Basalt
1173.8	17.4892	91.7189	91.7599	10.4932	1.77907	3.43654	4.26932	3.33156	-999.25	0.882429	2633.43	1936.98	1314.03	Basalt
1173.9	17.1544	91.7278	91.7599	10.4941	1.77903	3.42107	4.2338	3.33664	-999.25	0.93075	2637.47	1939.17	1313.36	Basalt
1174	16.8075	91.7367	91.7732	10.4951	1.77856	3.27169	4.05946	3.15978	-999.25	0.938567	2640.17	1936.98	1311.69	Basalt
1174.1	16.5386	91.7455	91.7865	10.4961	1.77884	3.1798	3.93971	2.87253	0.35022	0.88674	2640.17	1936.98	1311.02	Basalt
1174.2	16.3759	91.7544	91.7865	10.497	1.77822	3.24362	3.94243	2.59318	0.784671	0.777756	2642.88	1936.26	1309.03	Basalt
1174.3	16.3238	91.7632	91.7998	10.498	1.77861	3.36536	3.98757	2.33957	1.22743	0.634904	2646.95	1937.71	1307.7	Basalt
1174.4	16.3261	91.772	91.7998	10.4989	1.77831	3.39906	4.00217	2.04628	1.62752	0.490687	2651.03	1938.44	1306.37	Basalt
1174.5	16.3117	91.7809	91.8131	10.4999	1.77796	3.34284	4.01199	1.67485	1.93304	0.374038	2655.12	1938.44	1304.39	Basalt
1174.6	16.2533	91.7896	91.8264	10.5009	1.7779	3.30744	4.06929	1.25889	2.09744	0.306784	2660.6	1937.71	1303.4	Basalt
1174.7	16.1634	91.7984	91.8264	10.5018	1.7777	3.33302	4.16681	0.859109	2.10492	0.299917	2664.73	1940.63	1303.07	Basalt
1174.8	16.0027	91.8079	91.8397	10.5027	1.77776	3.34242	4.25737	0.522807	1.99254	0.34644	2668.87	1943.55	1303.07	Basalt
1174.9	15.7176	91.8183	91.853	10.5036	1.77734	3.29278	4.30934	0.324064	1.80505	0.428622	2671.63	1946.49	1301.1	Basalt
1175	15.3798	91.8288	91.853	10.5046	1.77743	3.2732	4.31487	0.360604	1.57428	0.520325	2673.02	1945.75	1301.1	Basalt
1175.1	15.1327	91.8381	91.8663	10.5056	1.77712	3.38722	4.29047	0.651454	1.33128	0.597919	2675.79	1947.22	1301.43	Basalt
1175.2	15.0825	91.8456	91.8795	10.5066	1.77713	3.59336	4.26046	1.08488	1.08811	0.658098	2682.75	1945.75	1299.46	Basalt
1175.3	15.2572	91.8521	91.8795	10.5075	1.77709	3.73423	4.21531	1.50134	0.857034	0.705957	2686.95	1946.49	1300.44	Basalt
1175.4	15.5992	91.8585	91.8929	10.5084	1.77691	3.72227	4.13748	1.79948	0.639116	0.753046	2688.35	1945.02	1300.11	Basalt
1175.5	15.9632	91.8656	91.9061	10.5094	1.77653	3.616	4.07008	1.9127	0.455415	0.806784	2689.75	1942.82	1299.79	Basalt

Figure 64 An example of well logging data in excel file from Mountain Home



REFERENCES

- Anders, M.H., et al. 2014. A fixed sublithospheric source for the late Neogene track of the Yellowstone hotspot: Implications of the Heise and Picabo volcanic fields. Journal of Geophysical Research: Solid Earth 119: 2871-2906.
- Bedrosian, P.A., and Feucht, D.W. 2014. Structure and tectonics of the northwestern United States from EarthScope USArray magnetotelluric data. Earth and Planetary Science Letters 402: 275-289.
- Bengio, Y., Simard, P., and Frasconi, P. 1994. Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks 5: 157-166.
- Berendsen, P., et al. 1988. Texaco Poersch 1, Washington County, Kansas: Preliminary Geologic Report of the Pre-Phanerozoic Rocks. Kansas Geological Survey.
- Bestagini, P., Lipari, V., and Tubaro, S. 2017. A machine learning approach to facies classification using well logs. SEG Technical Program Expanded Abstracts 2017, 2137-2142.
- Bishop, C.M. 2006. Pattern Recognition and Machine Learning. Singapore: Springer Science Business Media, LLC.
- Brand, B.D., and White, C.M. 2007. Origin and stratigraphy of phreatomagmatic deposits at the Pleistocene Sinker Butte Volcano, Western Snake River Plain, Idaho. Journal of Volcanology and Geothermal Research 160: 319-339.
- Breiman, L. 2001. Random Forests. Machine Learning 45: 5-32.
- Catalogna, M., Cohen, E., Fishman, S., Halpern, Z., Nevo, U., and Ben-Jacob, E. 2012. Artificial neural networks based controller for glucose monitoring during clamp test. PLoS One 7: 1-10.
- Chen, T., and Guestrin, C. 2016. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16, 785-794.
- Cristianini, N., and Taylor, S. 2000. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press.
- Cun, Y.L., et al. 1989. Handwritten digit recognition: applications of neural network chips

- and automatic learning. IEEE Communications Magazine 27: 41-46.
- DeCourten, F. 2009. Geology of Northern California. Cengage Learning.
- Devak, M., Dhanya, C.T., and Gosain, A.K. 2015. Dynamic coupling of support vector machine and K-nearest neighbour for downscaling daily rainfall. Journal of Hydrology 525: 286-301.
- Dubois, M.K., Bohling, G.C., and Chakrabarti, S. 2007. Comparison of four approaches to a rock facies classification problem. Computers & Geosciences 33: 599-617.
- Ellis, B.S., et al. 2010. Petrologic constraints on the development of a large-volume, high temperature, silicic magma system: The Twin Falls eruptive centre, central Snake River Plain. Lithos 120: 475-489.
- Engebretson, D.C., Cox, A., and Gordon, R.G. 1984. Relative motions between oceanic plates of the Pacific Basin. Geophysical Research Atmospheres 10291-10310.
- Fanchi, J.R. 2002. Well Logging. In Fanchi, J.R. (ed.), Shared Earth Modeling, pp. 52-68. Woburn: Butterworth-Heinemann.
- Friedl, M.A., and Brodley, C.E. 1997. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment 61: 399-409.
- Friedman, J.H. 2002. Stochastic gradient boosting. Computational Statistics & Data Analysis 38.
- Glowacz, A., and Glowacz, Z. 2016. Recognition of images of finger skin with application of histogram, image filtration and K-NN classifier. Biocybernetics and Biomedical Engineering 36: 95-101.
- Hall, M., and Hall, B. 2017. Distributed collaborative prediction: Results of the machine learning contest. The Leading Edge 36: 267-269.
- He, K., Zhang, X., Ren, S., and Sun, J. 2016. Deep Residual Learning for Image Recognition. in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770-778. 27-30 June 2016.
- Huang, W., Zhu, Y., and Xiao, L. 2015. Methods of phase division and distribution prediction of volcanic rocks in the Anda Sag, Xujiaweizi Fault Depression, Songliao Basin. Natural Gas Industry B 2: 307-313.
- Hughes, S.S., Smith, R.P., Hackett, W.R., and Anderson, S.R. 1999. Mafic volcanism and environmental geology of the eastern Snake River Plain. In Hughes, S.S. and

- Thackray, G.D. (eds.), Guidebook to the Geology of Eastern Idaho, pp. 143-168. Idaho: Idaho Museum of Natural History.
- Humphreys, E.D. 1995. Post-Laramide removal of the Farallon slab, western United States. Geology 23.
- Konaté, A.A., et al. 2015. Machine Learning Interpretation of Conventional Well Logs in Crystalline Rocks. International Conference on Swarm Intelligence: ICSI 2015, Cham. 360-370.
- Kuntz, M.A., Covington, H.R., and Schoor, L.J. 1992. An overview of basaltic volcanism of the Eastern Snake River Plain, Idaho. In Link, P.K., Kuntz, M.A., and Platt, L.B. (eds.), Regional Geology of Eastern Idaho and Western Wyoming, pp. 227-269. Geological Society of America Memoir 179.
- Lewis, R.S., Link, P.K., Stanford, L.R., and Long, S.P. 2012. The Geologic Map of Idaho. [Online]. Available from: www.idahogeology.org/product/m-9
- Lin, Z., Pan, H., Fang, H., Gao, W., and Liu, D. 2018. High-altitude well log evaluation of a permafrost gas hydrate reservoir in the Muli area of Qinghai, China. Sci Rep 8: 12596.
- Liu, C.R. 2017. Introduction to Well Logging. In Theory of Electromagnetic Well Logging, pp. 1-16.
- Liu, J., Wu, C., and Wang, J. 2018. Gated recurrent units based neural network for time heterogeneous feedback recommendation. Information Sciences 423: 50-65.
- Liu, Y., et al. 2017. Detecting Cancer Metastases on Gigapixel Pathology Images. ArXiv 1-13.
- Lonsdale, P. 2005. Creation of the Cocos and Nazca plates by fission of the Farallon plate. Tectonophysics 404: 237-264.
- Lopes, R., and Jorge, A. 2017. Mind the Gap: A Well Log Data Analysis. The Workshop on Data Mining for Oil and Gas, Houston, Texas, USA. 1-6.
- López, I., Aragonés, L., Villacampa, Y., and Serra, J.C. 2017. Neural network for determining the characteristic points of the bars. Ocean Engineering 136: 141-151.
- Martinius, A.W., Geel, C.R., and Arribas, J. 2002. Lithofacies characterization of fluvial sandstones from outcrop gamma-ray logs (Loranca Basin, Spain): the influence

- of provenance. Petroleum Geoscience 8: 51-62.
- Mazurek, J. 2004. Genetic controls on basalt alteration within the eastern Snake River Plain aquifer system. Master Degree, Geosciences Idaho, Idaho State University.
- Mcculloch, W.S., and Pitts, W. 1943. A logical Calculus of the ideas immanent in nervous activity. Bulletin of Mathematic Biophysics 5: 115-133.
- McLing, T.L., Smith, R.P., Smith, R.W., Blackwell, D.D., Roback, R.C., and Sondrup, A.J. 2016. Wellbore and groundwater temperature distribution eastern Snake River Plain, Idaho: Implications for groundwater flow and geothermal potential. Journal of Volcanology and Geothermal Research 320: 144-155.
- McQuarrie, N., and Rodgers, D.W. 1998. Subsidence of a volcanic basin by flexure and lower crustal flow: The eastern Snake River Plain, Idaho. Tectonics 17: 203-220.
- Meyer, D., Leisch, F., and Hornik, K. 2003. The support vector machine under test. Neurocomputing 55: 169-186.
- Neupane, G., Mattson, E.D., McLing, T.L., Palmer, C.D., Smith, R.W., and Wood, T.R. 2014. Deep Geothermal Reservoir Temperatures in the Eastern Snake River Plain, Idaho using Multicomponent Geothermometry. in Thirty-Ninth Workshop on Geothermal Reservoir Engineering. pp. 831-843. Stanford University, Stanford, California. Stanford Geothermal Program.
- Panday, D., Cordeiro de Amorim, R., and Lane, P. 2018. Feature weighting as a tool for unsupervised feature selection. Information Processing Letters 129: 44-52.
- Parsons, T., Thompson, G.A., and Smith, R.P. 1998. More than one way to stretch: a tectonic model for extension along the plume track of the Yellowstone hotspot and adjacent Basin and Range Province. Tectonics 17: 221-234.
- Perkins, M.E., Brown, F.H., Nash, W.P., Williams, S.K., and McIntosh, W. 1998. Sequence, age, and source of silicic fallout tuffs in middle to late Miocene basins of the northern Basin and Range province. Geological Society of America Bulletin 110: 344-360.
- Pierce, K.L., and Morgan, L.A. 1992. The Track of the Yellowstone Hotspot: Volcanism, faulting, and uplift. In Link, P.K., Kuntz, M.A., and Platt, L.B. (eds.), Regional Geology of Eastern Idaho and Western Wyoming, pp. 1-53. Geological Society of America Memoir 179.

- Pierce, K.L., and Morgan, L.A. 2009. Is the track of the Yellowstone hotspot driven by a deep mantle plume? — Review of volcanism, faulting, and uplift in light of new data. Journal of Volcanology and Geothermal Research 188: 1-25.
- Podgorney, R. 2016a. Snake River Plain FORGE Site Characterization Data. [Online]. 04/21/2016. Available from: <http://gdr.openet.org/submissions/793> [12/20/2017]
- Podgorney, R.K. 2016b. Phase 1 Topical Report Snake River Geothermal. Idaho Falls, Idaho: Idaho National Laboratory.
- Rastegarzadeh, G., and Nemati, M. 2018. Primary mass discrimination of high energy cosmic rays using PNN and k-NN methods. Advances in Space Research 61: 1181-1191.
- Rodgers, D.W., Ore, H.T., Bobo, R.T., McQuarrie, N., and Zentner, N. 2002. Extension and subsidence of the eastern Snake River Plain, Idaho. In Bonnicksen, B., White, C.M., and McCurry, M. (eds.), Tectonic and Magmatic Evolution of the Snake River Plain Volcanic Province, pp. 121-155. Idaho Geological Survey Bulletin 30.
- Roy, S. 2013. Near-surface Characterization via Seismic Surface-wave Inversion. PhD, Department of Earth and Atmospheric Sciences University of Houston.
- Shen, L., Cao, D., Xu, Q., Huang, X., Xiao, N., and Liang, Y. 2016. A novel local manifold-ranking based K-NN for modeling the regression between bioactivity and molecular descriptors. Chemometrics and Intelligent Laboratory Systems 151: 71-77.
- Shervais, J. 2014a. Kimama Well - Borehole Geophysics Database. [Online]. 02/11/2014. Available from: <https://gdr.openet.org/submissions/284> [12/20/2017]
- Shervais, J. 2014b. Kimberly Well - Borehole Geophysics Database. [Online]. 02/05/2014. Available from: <https://gdr.openet.org/submissions/284> [12/20/2017]
- Shervais, J. 2014c. Mountain Home Well - Borehole Geophysics Database. [Online]. 02/06/2014. Available from: <https://gdr.openet.org/submissions/284> [12/20/2017]
- Shervais, J.W., et al. 2013. First Results from HOTSPOT: The Snake River Plain Scientific Drilling Project, Idaho, U.S.A. Scientific Drilling 15: 36-45.
- Shervais, J.W., and Vetter, S.K. 2009. High-K alkali basalts of the Western Snake River Plain (Idaho): Abrupt transition from tholeiitic to mildly alkaline plume-derived basalts. Journal of Volcanology and Geothermal Research 188: 141-152.

- Simonyan, K., and Zisserman, A. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRR abs/1409.1556.
- Smirnoff, A., Boisvert, E., and Paradis, S.J. 2008. Support vector machine for 3D modelling from sparse geological information of various origins. Computers & Geosciences 34: 127-143.
- Soltani, S., Kordestani, M., and Karim Aghaee, P. 2016. New estimation methodologies for well logging problems via a combination of fuzzy Kalman filter and different smoothers. Journal of Petroleum Science and Engineering 145: 704-710.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research 15: 1929 - 1958.
- Steinbach, M., and Tan, P.-N. 2009. kNN: k-Nearest Neighbors. In Wu, X. and Kumar, V. (eds.), The Top Ten Algorithms in Data Mining, pp. 151-161. New York: Taylor & Francis Group, LLC.
- Szegedy, C., et al. 2015. Going deeper with convolutions. in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1-9. 7-12 June 2015.
- Tassell, J.V., Ferns, M., McConnell, V., and Smith, G.R. 2001. The mid-Pliocene Imbler fish fossils, Grande Ronde Valley, Union County, Oregon, and the connection between Lake Idaho and the Columbia River. Oregon Geology 63: 77-96.
- Tester, J.W., et al. 2006. Geothermal ResourceBase Assessment. In The future of geothermal energy: impact of Enhanced Geothermal Systems (EGS) on the United States in the 21st Century, pp. 35-81. Geothermal Program, MS 3830 Renewable Energy and Power Department Idaho National Laboratory: U.S. Department of Energy.
- Tiab, D., and Donaldson, E.C. 2012. Basic Well-Log Interpretation. In Petrophysics, pp. 803-827.
- Tsai, C.-F., Li, M.-L., and Lin, W.-C. 2018. A class center based approach for missing value imputation. Knowledge-Based Systems 151: 124-135.
- Tsangaratos, P., and Ilia, I. 2016. Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size. Catena 145: 164-179.

- Tutubalina, E., Miftahutdinov, Z., Nikolenko, S., and Malykh, V. 2018. Medical concept normalization in social media posts with recurrent neural networks. J Biomed Inform 84: 93-102.
- Wang, G., Cheng, G., and Carr, T.R. 2013. The application of improved NeuroEvolution of Augmenting Topologies neural network in Marcellus Shale lithofacies prediction. Computers & Geosciences 54: 50-65.
- Whitehead, R.L. 1992. Geohydrologic framework of the Snake River plain regional aquifer system, Idaho and eastern Oregon. U.S. Geological Survey Professional Paper 1400 - 1432.
- Whittemore, D.O., et al. 1997. The Dakota Aquifer Program Annual Report, FY92. Kansas Geological Survey.
- Wood, S.H., and Clemens, D.M. 2002. Geologic and tectonic history of the western Snake River Plain, Idaho and Oregon. In Bonnichsen, B., White, C.M., and McCurry, M. (eds.), Tectonic and Magmatic Evolution of the Snake River Plain Volcanic Province, pp. 69-103. Idaho Geological Survey Bulletin.
- Wu, H., and Zhao, J. 2018. Deep convolutional neural network model based chemical process fault diagnosis. Computers & Chemical Engineering 115: 185-197.
- Xie, Y., Zhu, C., Zhou, W., Li, Z., Liu, X., and Tu, M. 2018. Evaluation of machine learning methods for formation lithology identification: A comparison of tuning processes and model performances. Journal of Petroleum Science and Engineering 160: 182-193.
- Xu, Q.S., and Liang, Y.Z. 2001. Monte Carlo cross validation. Chemometrics and Intelligent Laboratory Systems 56: 1-11.
- Xue, H., Yang, Q., and Chen, S. 2009. SVM: Support Vector Machines. In Wu, X. and Kumar, V. (eds.), The Top Ten Algorithms in Data Mining, pp. 37-59. New York: Taylor & Francis Group, LLC.
- Yadav, N., Yadav, A., and Kumar, M. 2015. An Introduction to Neural Network Methods for Differential Equations. New York: Springer.
- Zhu, L., Zhang, C., Wei, Y., and Zhang, C.-m. 2016. Permeability Prediction of the Tight Sandstone Reservoirs Using Hybrid Intelligent Algorithm and Nuclear Magnetic Resonance Logging Data. Arabian Journal for Science and Engineering 42.

Zhu, L., et al. 2018. Prediction of Total Organic Carbon Content in shale reservoir Based on a New Integrated Hybrid Neural Network and Conventional Well Logging Curves. Journal of Geophysics and Engineering 15: 1050-1061.





จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

VITA

NAME Worapop Thongsame

DATE OF BIRTH 08 May 1995

PLACE OF BIRTH Ratchaburi

INSTITUTIONS ATTENDED Chulalongkorn University

HOME ADDRESS 19/1 Khwaeng Thung Phaya Thai, Phaya Thai, Bangkok
10400, Thailand

PUBLICATION -

AWARD RECEIVED -

