Machine Learning for Natural Resource Assessment: An Application to the Blind Geothermal Systems of Nevada

Stephen Brown[1,2], Mark Coolbaugh[3], Jacob DeAngelo[4], James Faulds[3], Michael Fehler[1], Chen Gu[1], John Queen[5], Sven Treitel[5], Connor Smith[3], and Eli Mlawsky[3]

[1]Massachusetts Institute of Technology, Cambridge, Massachusetts; [2]Aprovechar Lab L3C, Montpelier, Vermont; [3]Great Basin Center for Geothermal Energy, Nevada Bureau of Mines and Geology, University of Nevada, Reno, Nevada; [4]U.S. Geological Survey, MoffettField, California; [5]Hi-Q Geophysical, Inc., Ponca City, Oklahoma

Keywords

Machine learning, play fairway analysis, Nevada, categorical features, structural control

ABSTRACT

A study is underway to apply machine learning methods to evaluate natural resource potential. In particular, we are considering the search for blind geothermal systems in Nevada. Beginning with the data and experience from the previous Nevada play fairway analysis project, we are building models in TensorFlow/Keras and gaining experience toward predicting the geothermal resource potential as a probability map. During the first year of this project we have encountered several issues particular to using geological and geophysical data sets with these tools. Through an illustrative example we develop a promising workflow for future use as more data become available and are analyzed.

1. Introduction

The Great Basin region is a world-class geothermal province with ~720 MWe of current power generation from ~24 power plants. Studies indicate far greater potential for both conventional hydrothermal and EGS systems in the region (Williams et al., 2009).

Because most geothermal systems in the Great Basin are controlled by Quaternary normal faults, they generally reside near the margins of actively subsiding basins. Thus, upwelling fluids along faults commonly flow into permeable subsurface sediments in the basin and do not reach the surface directly along the fault. Outflow from these upwellings may emanate many kilometers away from the deeper source or remain *blind* with no surface manifestations (Richards and Blackwell, 2002). Blind systems are thought to comprise the majority of geothermal resources in the region (Coolbaugh et al., 2007). Thus, techniques are needed to both identify the structural settings enhancing permeability and to determine which areas may host subsurface hydrothermal fluid flow.

Geothermal play fairway analysis (PFA) is a concept adapted from the petroleum industry (Doust, 2010) that aims to improve the efficiency and success rate of geothermal exploration and drilling by integration of geologic, geophysical, and geochemical parameters indicative of geothermal activity. A prior demonstration project (Faulds et al., 2017) focused on defining geothermal play fairways, generating detailed geothermal potential maps for ~1/3 of Nevada, and facilitating discovery of blind geothermal fields. The Nevada play fairway project incorporated ~10 geologic, geophysical, and geochemical parameters indicative of geothermal activity. It led to discovery of two new geothermal systems. The PFA leveraged logistic regression, weights of evidence, and other statistical measures as a type of machine learning technique. A set of features, each gauged by a perceived weight of influence, were combined to estimate geothermal potential. However, key limitations and challenges affected the PFA, including estimating weights of influence for parameters, incomplete data sets, and a limited number of training sites. We are now building upon the original Nevada PFA study of Faulds et al. (2017) to include principles and techniques of machine learning (Faulds et al., 2020). Here, we present some early results.

2. A Perspective on Machine Learning

Many problems in Earth sciences concerning the environment, energy, and natural hazards involve taking a set of observations on a map and making inferences about some unknown or unseen characteristic: such as predicting a numerical quantity or probability, recognizing a particular structure or category, or to build a new classification or taxonomy. Applications of practical importance are energy and mineral prospecting and resource assessment, location of buried objects of specific types such as land mines and tunnels, prediction of pathways for migration of groundwater contaminants, or to search for anomalies within a complex background as in security monitoring, to name a few.

To tackle these problems, one is tempted to bring to bear the recent developments and successes in artificial intelligence (AI) and its popular implementations known as machine learning, especially artificial neural networks including either fully-connected or convolutional deep neural networks (e.g., Goodfellow et al., 2016). However, the nature of the data for many Earth science applications presents a set of special considerations making direct application of many successful AI methods difficult.

First, our problems generally have a paucity of examples with known labels (training sites) for problems for which we would hope to apply supervised learning methods. Many problems in computer vision such as facial recognition, autonomous driving image recognition, etc., have many tens to perhaps hundreds of thousands of labeled examples from which to train, develop, and test deep network algorithms and architectures. While some geoscience problems such as hyper-spectral satellite imaging have this much training data (e.g., Xu et al., 2017), many other geoscience problems may have only on the order of one hundred or fewer labeled examples.

Another consideration is the nature of the data itself. Input features carrying important information for the regression or classification will necessarily be a mixture of numerical values (real numbers such as temperature, distance from a geologic fault, or gravity anomaly), categorical variables (mineral assemblages or rock type descriptions), and ordinal variables (i.e., ranked as this is bigger than that, but with no scaling). Geologic data will commonly be indexed to maps and thus are related spatially to other features. These data, however, may not have the same resolution nor the same degree of certainty.

Finally, there are physical principles, variably understood ahead of time, which point to reasonable relationships between features and labels. Incorporating such “expert” knowledge into the hypotheses is important, especially to counter the problem of a small number of examples and to ensure proper weighting of the uncertainty of various data sources.

Generally, supervised machine learning uses two sources of knowledge: (1) labeled data (x,y) pairs and (2) hand-engineered features, network architecture, and other components. For the cases of “small data” (relatively few labeled feature-outcome pairs) we often use more hand-engineering and/or specialized algorithms. For the cases of “big data” we can generally apply simpler algorithms and less hand-engineering. For small data, hand engineering is commonly the most efficient way to better learning performance.

One would like to strike a balance between the ideal of feeding large numbers of features into a network and letting the algorithm determine the relationships and, on the other extreme, engineering the complete hypothesis and algorithm by hand. The former, while being unbiased and allowing the data to speak for themselves, can be prone to overfitting and for this application is perhaps at best inefficient or at worst intractable. The latter runs the risk of extreme bias leading to under-fitting such that important links among features will never be recognized. Also, the algorithms developed in the hand-engineered approach may not be extensible to new data types nor to new areas of application.

We propose to address these issues by taking advantage of the previous Nevada geothermal PFA study. Through revisiting this work and re-analyzing the accompanying data set using machine learning methods, we can gain experience in bringing deep learning principles to application on this important class of Earth science problems.

3. Starting with the PFA Workflow and Results

The original Nevada geothermal play fairway analysis as applied to the Great Basin resulted in the workflow illustrated in Figure 1. Information gathered and quantified throughout the study area (listed in the colored boxes) are combined with information from other sources as you follow the arrows through the diagram. Specific geological and geophysical features were chosen as they are thought to indicate essential aspects of permeability and heat sources. First, these are combined algebraically as shown to calculate the “fairway,” a number scaled to indicate the degree of the geothermal potential. Then, this fairway map is combined with further evidence and constraints to define a map of the “Exploration Opportunities.” The combination of terms is done by multiplying each parameter by a unique “weight” and then adding this product to the result of other calculations shown in the workflow. The weights used in this analysis were derived using a combination of statistics, including Bayesian-based weights-of-evidence and logistic regression (e.g., red numbers) through analysis of known benchmark sites in the study area, and expert judgment (black numbers) due to known limitations of some data sets and small number of training sites. In this paper we consider only the workflow up to the point of the fairway.

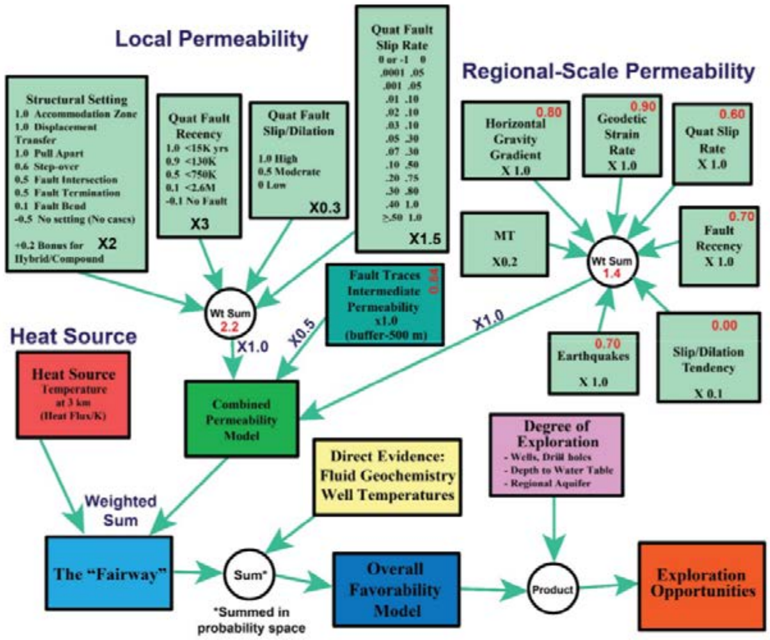


Figure 1: Nevada PFA workflow. Note the mixture of numerical and categorical/ordinal features, each tied to geographic positions on a map with varying scales of resolution (modified from Faulds et al., 2017).

Following this statistically based, expert-assisted analysis and constrained by benchmark sites, the fairway map of the study area within the Great Basin region was produced (Figure 2). This map represents the geothermal potential. Subsequent drilling in two promising areas on the map revealed viable geothermal resources (Faulds et al., 2018, 2019; Craig, 2018), so the method of geothermal play fairway analysis is thought to bear some practical predictive power.

Encouraged by this success, we strive in the current study to reevaluate this problem using machine learning techniques and use this map as a point of reference for our results.

4. Define a Supervised Learning Problem

If we redraw the original PFA workflow diagram (Figure 1), we immediately recognize that this computation graph is a relatively simple feed-forward network (Figure 3a), in effect a highly-engineered neural network where the nodes are connected in a very specific way. Linear combinations of feature weights are linked through linear "activations" to the next layer in the computation. This is a feed-forward network since we go directly from the measured features on the left through the series of algebraic calculations, without any loops, as we move toward the resulting fairway value on the right.

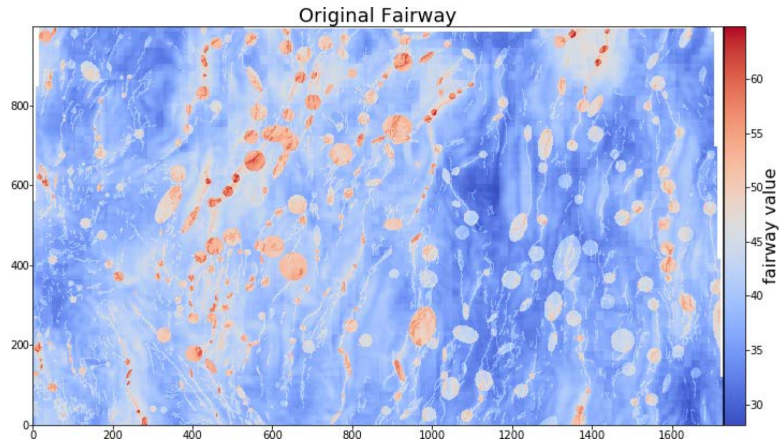


Figure 2: The play fairway model of the study area in west-central to eastern Nevada (modified from Faulds et al., 2018). This map represents the geothermal potential, or "fairway value," in a numeric scale from low to high indicated by cool to warm colors. On this map and the others which follow, the x and y axis labels represent the indices of the 250 by 250 meter grid blocks dividing up the study area.

We can see, then, the potential for replacing the statistically-based and expert-assisted methods for determining the model parameters (weights and biases) with machine learning optimization methods. This is done by example and is known as a "supervised" learning problem where we "learn" a mapping technique from "features" (geological and geophysical properties) to "labels" (whether a site has geothermal potential or not). The weights in the network are determined through an automatic numerical optimization procedure, by repeatedly presenting the algorithm with a labeled training set of benchmark examples with known combinations of features and labels. Once optimized, the final network can then be used to predict the probability of a potential geothermal system by feeding into it a new set of features. Thus, a broader predictive map of geothermal potential can be produced.

The machine learning approach has some advantages. First, the outcome can be cast as a probability, defensible through validation tests. Second, careful implementation can reduce or eliminate biases in the choices of the most appropriate features and in the choices of the network architecture controlling how the features are combined. Finally, the algorithms can be easily automated, generalized and refined, and extended to accommodate new data sources.

In our new work we begin with the network defined previously (Figure 3a) and move to more complex and general fully-connected neural networks (Figure 3b). We use the data sets that the previous project provides. As can be seen in the original workflow diagram, we have sets of data divided by their perceived information content indicating local, intermediate, and regional scale permeability and heat. For example, the local permeability is characterized by four features: (1) structural setting, (2) Quaternary fault recency, (3) Quaternary fault slip and dilation tendency, and (4) Quaternary fault slip rate. These parameters are known in map form throughout the study area and are referred to as "features." At certain geographic locations there are benchmarks, where it is known whether or not a geothermal potential exists, i.e., an existing power plant or definitive positive or negative drilling results. These comprise the "labels" referred to as positive or negative sites. The features and label pairs at the benchmark sites comprise the training and test (or validation) data sets for the analysis, and the goal is to predict the label as a probability everywhere else in the region where the features are known.

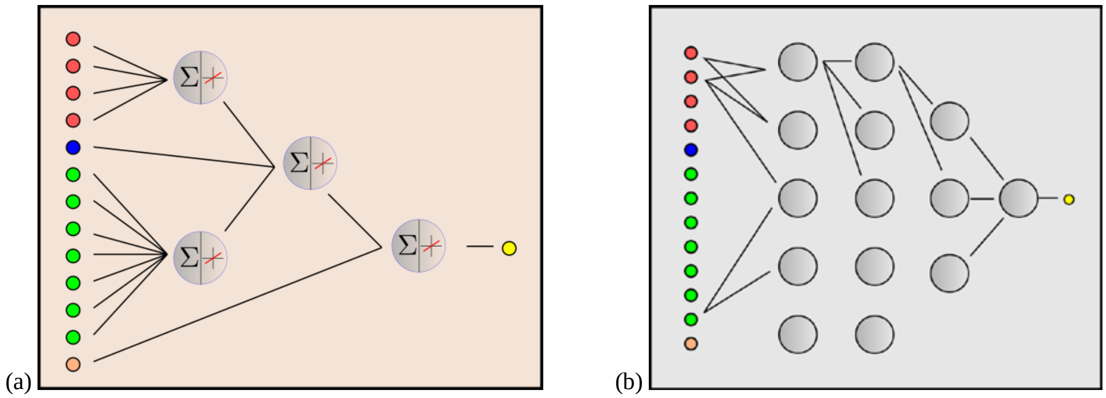


Figure 3: (a) The PFA workflow redrawn with the features on the left and the fairway on the right. For visual reference the various related feature categories on the left representing the scales of permeability and heat sources are colored. (b) A general fully-connected neural network configured to take the PFA features as inputs on the left and infer the geothermal resource potential as a probability on the right. Note that this diagram is schematic only and for simplicity does not show all possible connections, layers, or neurons. Refer to Goodfellow et al. (2016) for a hint to the many possibilities.

5. A Look at the Data Sets

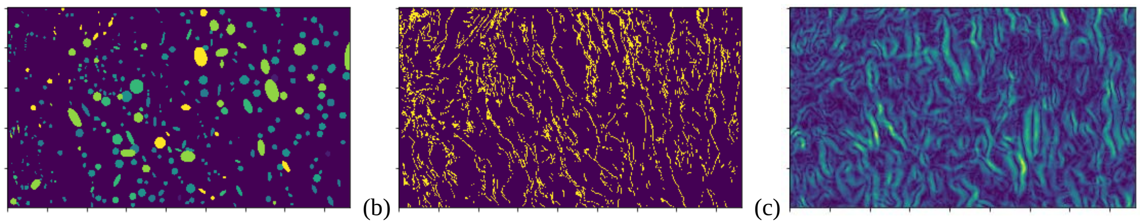
As the machine learning methods we apply are fundamentally algebraic computations, it is important to take a closer look at the data we have on hand. The data have been divided into 4 groups in the original workflow as determined by their information content. Some of these data are numerical and others are categorical and divisions among them fall into the three distinct cases illustrated in Figure 4.

The easiest to understand are the continuous numerical features. All geophysical maps pertaining to regional permeability and the heat source fall into this category. These maps represent continuous real numbers known at every grid block in the study area. For example Figure 4c shows the horizontal gravity gradient.

The next feature type is binary (two categories). The intermediate permeability is represented by a line drawing of fault traces (Figure 4b) where numerically it has value 1 where a fault exists and 0 otherwise. While technically a categorical feature (1’s and 0’s), this map is populated quite densely and uniformly in the study area, and thus it could be meaningfully transformed into a numerical feature in one of several ways as we will illustrate later.

Finally, there are the categorical features representing the local permeability (Figure 4a shows the structural settings). These maps have two distinctive characteristics. First there are finite numbers of categories involved for each. For example, there are 8 structural settings defined with an additional qualifier. These categories have further been assigned a number by the experts ranking their perceived importance to defining their influence on the fairway. Second, these features are defined only within ellipsoidal regions fairly heterogeneously and sparsely within the study area, whereas in between indicates either "no importance" or "unknown."

As we will discuss further, to use these three disparate types of features effectively together in a machine learning algorithm, we must place them on an equal footing numerically by either pre-processing or other considerations.



(a) (b) (c)

Figure 4: Examples of feature types encountered in this problem. (a) The map of favorable structural settings as an example of the categorical features. (b) The map of fault traces as an example of a dense binary feature. (c) The map of horizontal gravity gradient as an example of a continuous numerical feature.

6. Issues to Confront

Now that we have defined a supervised learning problem to solve and have taken a closer look at the available data, we have been working toward producing a probabilistic equivalent of the fairway map to compare to the outcome of the original PFA project. This has involved translating the data set into a form usable by machine learning algorithms and implementing those algorithms in TensorFlow/Keras (Abadi et al., 2016). In developing these new models and investigating their performance and utility, we have had to confront and overcome several issues in applying machine learning to this type of resource assessment problem. As this is a work in progress, we simply list them here and then follow with one illustration leading to a deeper understanding. As we will see, this leads to a promising workflow for continued work.

- <u>Small numbers of examples (initially only 34 positive benchmarks).</u> Generally, even a simple neural network as used in the original PFA study (Figure 3a) or a minimal fully-connected neural network (e.g., Figure 3b) are capable of fitting quite a large range of nonlinear multivariate relationships between the features and the labels. Small numbers of examples can lead to overfitting (the data are fit almost exactly) yet have little capability for generalization (being able to accurately predict the labels of data sets not used in the training process). Remedies for this problem are as follows: (1) acquire more examples; (2) use some regularization methods to prevent overfitting by limiting the size or importance of any one weight in the network over others; (3) use a method called "dropout" to prevent overfitting by randomly omitting different links in the network during each training cycle and later restoring them all during inference to make the trained network not too dependent on one type of feature or parameter; or (4) use a transfer learning method whereby a network successfully trained on a similar task or trained on another data set is re-trained or fine-tuned on a new problem, thus transferring previously learned skills.

- Potential imbalance in the training data set. This refers to unequal numbers of examples in each of the classes (positive and negative geothermal sites). If too many examples are presented to the algorithm of one class over the other, the network can become biased by naively using standard optimization techniques. The extreme case is, for example, if only positive examples are given. Then, the algorithm can achieve near 100% accuracy on the training set by simply always outputting a positive result regardless of evidence from the feature set. Remedies for this problem are acquiring more data, trimming the data set, or data augmentation or simulation.

- A mixture of categorical and continuous numerical features. The neural networks ultimately use numerical computations internally for discriminating among the classes. Categorical features, such as geologic concepts and descriptions need to be replaced by some kind of numbering scheme. One possibility is to rank the categories by importance in a numerical scale. This is fine if the assignment follows a natural order or is justified through some physics. Another possibility is to use a numerical encoding method where the categories are nominal, set as distinct from one another, yet one category is not preferred over another. The advantage here is that evidence is presented in the training phase, and the algorithm determines the importance of the features during optimization.

- Data are not uniformly distributed or are partly unknown. Missing data is a recognized problem in machine learning, since the algorithms expect values to be present in each part of the numerical arrays. If there are missing data, then a substitute must be made to keep the computations functioning properly. Missing data can be imputed by some procedure like averaging nearby points, inserting a constant (e.g., 0.0) or filling with a random number sampled from some appropriate distribution. In any case, the network predictions will be less dependent on values of those features that are occasionally missing.

- An appropriate neural network architecture and parameters must be chosen. We desire to construct a neural network general enough and capable enough to model our data and generalize well to new situations. Thus, we at a minimum must extend the PFA workflow from its original form of Figure 3a to a fully-connected network as in Figure 3b. In doing this we want to minimize overfitting, and we want as few parameters as possible to make the computations efficient. The choice of network architecture is very much an art, yet it can be quantified. Given metrics of network suitability (how easily trained, how well they fit the training set, how well they generalize, and their degree of complexity), we can use searching methods, such as genetic algorithms, to find the best candidates for the job.

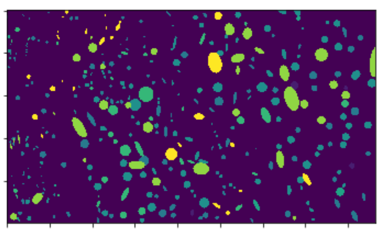
In the course of our work so far on applying machine learning methods to geothermal resource assessment, we have encountered each of these issues and have explored remedies for them. In what follows we highlight one of these problems and show how this leads to a promising workflow for the future.

7. Example Problem: Use of Categorical Data

One issue we have encountered in our work is how to consider the use of mixed continuous numerical data and sparsely known categorical data. Of main concern are the local permeability features, such as the structural settings, known only in elliptical regions throughout the study area (Figure 4a). We also have the problem of how to best treat the binary features that represent the fault traces (Figure 4b).

The favorable structural settings category in the local permeability feature maps (Figure 5a) were assigned a numerical scale in terms of importance on the basis of analysis of the benchmarks in the study area and power-producing systems throughout the Great Basin region (e.g., Faulds et al., 2013, 2015). This could lead to some bias in the outcome if the numerical scale is not completely correct. This possibility can be checked by performing two sets of simulations: one using the assigned scale to prescribe the importance of each category (Figure 5b) or by using a one-hot binary encoding for each category which treats the categories as distinct yet of potentially equal importance (Figure 5c). This second possibility will result in the optimization algorithm later weighting the importance of each category appropriately based on the training evidence provided.

(a)



(b)

```
1.0 Accommodation
1.0 Disp Transfer
1.0 Pull Apart
0.6 Step Over
0.5 Intersection
0.5 Termination
0.1 Bend
-0.5 None
```

(c)

```
[1 0 0 0 0 0 0 0]
[0 1 0 0 0 0 0 0]
[0 0 1 0 0 0 0 0]
[0 0 0 1 0 0 0 0]
[0 0 0 0 1 0 0 0]
[0 0 0 0 0 1 0 0]
[0 0 0 0 0 0 1 0]
[0 0 0 0 0 0 0 1]
```

Figure 5: Ways to use categorical features in a neural network. (a) The map of structural settings, where categories are labeled only within the ellipses. (b) These categories have been given a numerical scale signifying their importance by statistical and expert analysis. This numerical assignment could impart a bias on the solution. (c) Alternatively, the categories can be represented in a binary "one-hot" encoding scheme to make them distinct, yet removing the apriori assignment of significance.

The binary features representing the Quaternary fault traces (Figure 6a) present a minor problem to be addressed. This map represents not only the fault traces, but also a 500 m buffer was drawn around these faults, as this would be within the window of not detecting a smaller favorable setting due to the lack of detailed mapping. Even so, the map is binary with 1's on and near the faults and 0's elsewhere, creating a data set that does not easily mix with the other numerical features. The map is dense enough that it may be fine to leave the data in their original form, yet it may be better to transform this binary feature map somewhat. One possibility is to compute an aerial fault density map by a smoothing process (e.g., Figure 6b) or another is to use the local distance to the nearest fault (e.g., Figure 6c). These latter two cases convert the binary map into a numerical one, completely compatible with the other geophysical maps. In our preliminary studies we have been working with the fault density variation.

A final, somewhat more difficult problem remains with the categorical features (favorable structural settings) of Figure 5. These are thought to be extremely important features for discrimination yet they are also heterogeneously distributed throughout the study area. In fact, all positive examples exist within the ellipses where these features are known, and most negative examples fall outside of these ellipses. This requires us to take further special care in mixing the data types. Cluster analysis shows that due to the heterogeneity of the feature sets the positive and negative sites are seen to be widely separated and it is too easy for a naive machine learning algorithm to divide them. As a consequence, we have found that direct use of the raw category values in the standard machine learning algorithms leads to suspect results and could be thought of as an overfitting problem. We illustrate this in the following example.

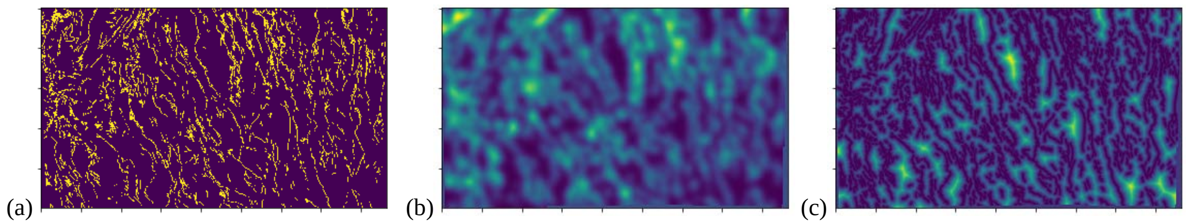


Figure 6: Ways to use the binary features. The map of fault traces (a) is a matrix of 1's and 0's indicating the hand-drawn sketch of fault patterns. These can be transformed to continuous numerical features easily by, for example, calculating the fault density (b) or by calculating the distance to the nearest fault (c).

To begin with, we know that the original PFA workflow (Figures 1 and 3a) faithfully produces the fairway map (Figure 2). Also, by considering almost any relatively simple fully-connected neural network (Figure 3b, with say 3 hidden layers of 64 neurons each for example) a set of weights in the network can be found to create an almost identical map. However, if we take that very same fully-connected network and train it on the PFA benchmarks, the optimized network produces maps similar to Figure 7a. It seems that there is a tendency for extreme over-fitting or becoming stuck in a local minimum in the optimization of the network, such that only where the categorical features are specified a prediction is available. This map is not very useful. If we do this again, but this time ignore all of the categorical features (by setting them to 0.0) then we produce the map of figure 7b. Here, there is lower predictive accuracy everywhere and too much of the study area appears to have high resource potential. Again, not a very useful result. We need to find a middle ground between these two extremes.

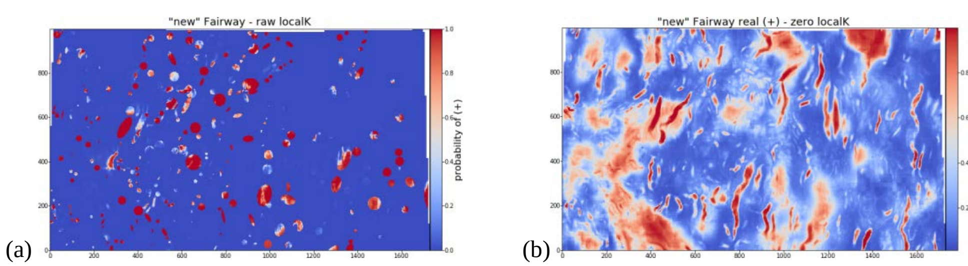
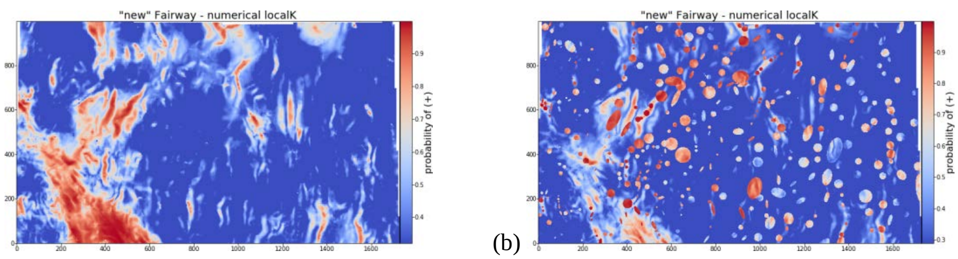


Figure 7: (a) Example of a fairway probability map resulting from optimization of a fully-connected neural network as in Figure 3b using the categorical features directly. This situation tends to over-fit the ellipsoidal regions producing a map of minimal use. (b) Example of a fairway probability map resulting from optimizing the same network, but in the training and inference the categorical features were ignored by setting them to zero everywhere. Here, the map indicates too large an area with high geothermal potential probability. Again, this map is of minimal use.

We have found the following solution. First, we augment our training data set by considering not just the benchmark sites themselves but also an aura of grid blocks around each one creating a much larger training set representing a kind of blurred view of the geothermal system. We train the network on these simulated data and ensure the training results in adequate accuracy on both training and validation data sets. Importantly, we also ensure that the categorical features are deemphasized in their importance during this training. This is done through a combination of regularization and dropout for these features. Once trained in this way, we might consider this network to be relatively good at describing the characteristics of a generic geothermal system like the Great Basin region. An example map produced at this stage is shown in Figure 8a, where a prediction accuracy for both training and previously unseen validation examples of around 70% has been obtained. This pre-trained network will now be used for transfer learning (e.g., Goodfellow et al., 2016).

It has been found that for many image classification problems based on convolutional neural networks, a network trained on one set of image classes (say species of animals) using many tens of thousands of training examples becomes good at discriminating a set of new classes after a fine-tuning step with many fewer examples. Thus, the experience gained in the initial training is transferred to the new task quite easily.

We take advantage of this idea here, as now we fine tune our generic network with only our benchmark sites and this time we allow the categorical features to retain their full significance. This results in maps like Figure 8b, showing not only >90% accuracy on both training and validation examples but with many of the salient features previously determined by the statistically based expert system in the original PFA project (Figure 2).



(a) (b)

Figure 8: (a) Example of a fairway probability map resulting from optimizing the same fully-connected neural network as in Figure 3b, but using a larger synthetic data set derived from the statistics of the original PFA benchmarks and at the same time de-emphasizing the categorical features. This model becomes our pre-trained master network to be used for transfer learning. (b) Example of the pre-trained master network fine-tuned using only the actual benchmark examples and using all numerical and categorical features. This example of transfer learning results in a fairway probability map containing many of the same features and areal coverage as the original fairway map of Figure 2.

8. Conclusion: A Proposed Workflow

The example of using categorical data alongside numerical data just discussed suggests a promising workflow use in applying machine learning to geothermal resource evaluation. We now have new training data available for the Nevada study area, and we will begin its analysis in the following way:

1. Augment real positive and negative training sites by using benchmark neighbors on the map.

2. Use genetic algorithms to find the “best” starting model architecture and hyper-parameters (e.g., Poli et al. 2008; Fortin et al., 2012).
3. Use the best model as the basis for data augmentation to create large simulated data sets for transfer learning. The data augmentation will involve both generative adversarial networks (GAN) (e.g., Goodfellow et al., 2016) and teacher-student networks (Xie et al., 2019).
4. Pre-train the best model on the large augmented data set with de-emphasis of categorical features.
5. Fine tune the network on all real training sites and all feature sets within our study area.

This process allows us to take advantage of transfer learning and honor the discriminating power of the categorical features without undue over-fitting. Data augmentation begins here with blurring our study area and then building upon this by generating simulated data honoring the statistical distributions for benchmarks we have on hand. In the future, however, we can imagine building our initial training data and our master network by considering data sets from other regions of the world ensuring that the examples come from like tectonic environments.

Acknowledgement

This material is based upon work supported by the U.S. Department of Energy’s Office of Energy Efficiency and Renewable Energy (EERE) under the Geothermal Technologies Program and Machine Learning initiative award number DE-EE0008762.

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X., 2016, TensorFlow: Large-scale machine learning on heterogeneous systems, In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), pp. 265-283.

Coolbaugh, M.F., Raines, G.L., and Zehner, R.E., 2007, Assessment of exploration bias in data-driven predictive models and the estimation of undiscovered resources: Natural Resources Research, v. 16, no. 2, p. 199-207.

Craig, Jason W., 2018, Discovery and analysis of a blind geothermal system in southeastern Gabbs Valley, western Nevada [M.S. Thesis]: University of Nevada, Reno, 111 p.

Doust, H., 2010, The exploration play: What do we mean by it?: American Association of Petroleum Geologists Bulletin, v. 94, p. 1657-1672.

Faulds, J.E., Hinz, N.H., Dering, G.M., Drew, D.L., 2013, The hybrid model – the most accommodating structural setting for geothermal power generation in the Great Basin, western USA: Geothermal Resources Council Transactions, v. 37, p. 3-10.

Faulds, J.E., Hinz, N.H., Coolbaugh, M.F., Siler, D.L., Shevenell, L.A., Queen, J.H., dePolo, C.M., Hammond, W.C., and Kreemer, C., 2015, Discovering blind geothermal systems in the Great Basin region: an integrated geologic and geophysical approach for establishing geothermal play fairways: Final report submitted to the Department of Energy, 245 p.

Faulds, J.E., Hinz, N.H., Coolbaugh, M.F., Shevenell, L.A., Sadowski, A.J., Shevenell, L.A., McConville, E., Craig, J., Sladek, C., and Siler D.L., 2017, Progress report on the Nevada play fairway project: Integrated geological, geochemical, and geophysical analyses of possible new geothermal systems in the Great Basin region: Proceedings, 42nd Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California, February 13-15, 2017, SGP-TR-212, 11 p.

Faulds, J.E., Craig, J.W., Coolbaugh, M.F., Hinz, N.H., Glen, J.M., Deoreo, S., 2018, Searching for blind geothermal systems utilizing play fairway analysis, western Nevada: Geothermal Resources Council Bulletin, v. 47, p. 34-42.

Faulds, J.E., Hinz, N.H., Coolbaugh, M.F., Ramelli, A., Glen, J.M, Ayling, B.A., Wannamaker, P.E., Deoreo, S.B., Siler, D.L., and Craig, J.W., 2019, Vectoring into potential blind geothermal systems in the Granite Springs Valley area, western Nevada: Application of the play fairway analysis at multiple scales: Proceedings 44th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California, SGP-TR-214, p. 74-84.

Faulds, J.E., Brown, S. Coolbaugh, M., DeAngelo, J., Queen, J.H., Treitel, s., Fehler, M., Mlawsky, E., Glen, J.M., Lindsey, C., Burns, E., Smith, C.M., Gu, C., and Ayling, B.F., 2020, Preliminary Report on Applications of Machine Learning Techniques to the Nevada Geothermal Play Fairway Analysis, Proceedings, 45th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California, February 10-12, 2020 SGP-TR-216.

Fortin, F., de Rainville, F., Gardner, M., Parizeau, M., and Gagné, C., 2012, DEAP: Evolutionary Algorithms Made Easy, Journal of Machine Learning Research, vol. 13, pp. 2171-2175.

Goodfellow, I., Bengio, Y., and Courville, A., 2016, Deep Learning, MIT Press.

Poli, R., Langdon, W.B., and McPhee, N.F., 2008. A Field Guide to Genetic Programming. Lulu Enterprises, UK Ltd.

Richards, M., and Blackwell, D., 2002, A difficult search: Why Basin and Range systems are hard to find: Geothermal Resources Council Bulletin, v. 31, p. 143-146.

Williams, C.F., Reed, M.J., DeAngelo, J., and Galanis, S.P. Jr., 2009, Quantifying the undiscovered geothermal resources of the United States: Geothermal Resources Council Transactions, v. 33, p. 995-1002.

Xie, Q., Hovy, E., Luong, M.T. and Le, Q.V., 2019. Self-training with noisy student improves ImageNet classification. arXiv preprint arXiv:1911.04252.

Zhu, X.X., et al., Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources, IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 4, pp. 8-36, Dec. 2017. doi: 10.1109/MGRS.2017.2762307.