ECEN765 Machine Learning with Networks

Course Project Proposal

Name: Liuyi Jin 225009797

Class: Application

**Introduction and Motivation**

In the petroleum engineering industry, some researchers are working on how to accurately predict the production of newly drilled wells according to the analysis of earlier production from older wells. In the past, when people try to predict production of conventional oil and gas reservoirs, they adopted deterministic prediction models. Those methods are derived from a fundamental equation in fluid mechanics area called Darcy’s Law. However, since 2008, unconventional reservoirs, which cannot be accurately predicted directly by using previous deterministic methods, are becoming important in the world energy market. Researchers then began to find innovative ways to accurately predict this kind of new unconventional reservoirs. Soon people in this area found that prediction methods with statistics principles underlined are very suitable in this particular occasion. Type well method is one of those methods.

Type wells, which are typical well production profiles based on analysis of existing well histories, are growing in acceptance in the industry as a mean of forecasting production in low-permeability reservoirs. The core idea of constructing type wells is to construct a well representing a set of wells that are being recognized as “analogous.” It aims to extract the inherent characteristics of multiple wells in a certain geologic area by producing one or a family of constructed wells which should be representative enough to represent all wells in an area of interest. The approach commonly employed in industry to construct type wells is to arithmetically average the production histories of a set of producing wells in the field. I will also take this method in this particular project.

Artificial Neural Networks, which are currently gaining much acceptance in the world. More and more people tends to employ this method within their own researches. There are several mathematical algorithms to train the neural networks. For example, steepest gradient descent, conjugate gradient descent, stochastic gradient descent (with momentum) are included in popular algorithms to train the target neural networks.

Our project is now aiming to train the given data set with given type wells, that is to say, we are going to classify the wells into several types using neural networks.

**Deliverable**

Given a large set of data (in our case, they are 200 production wells production). We already know that they have been labelled with corresponding class label (type I, type II, … ). The number of types is finite. Our goal is to train a neural network with a subset of the original 200 data samples. Here we select a subset of those samples because we need to save the rest samples to validate our training results. The expected result is that the trained network with its weights updated, which can be used to correctly classify undrilled wells. As for the algorithm used in this project, it would be stochastic gradient descent. The process can be graphically represented as shown in figure 1. The figure below did not show the back propagation.

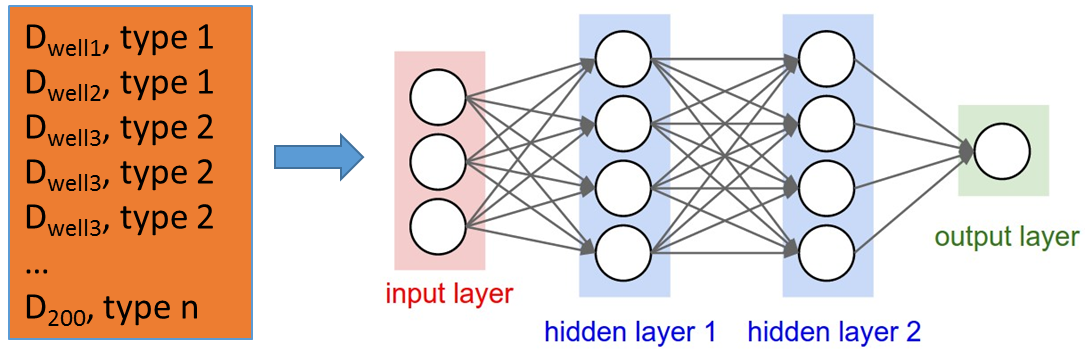


Figure 1 Schematic Representation of Type Well Data Training (http://cs231n.github.io/neural-networks-1/)

**Methodology**

After the type wells in a particular geologic area are constructed, they are claimed to be representative enough with a certain confidence. We train the rest data to fit them into one of those type wells. We will choose artificial neural network as our primary training method to classify newly drilled wells into one of our type wells. We would also try different other machine learning methods including KNN, Naïve Bayesian, etc., for comparison convenience. If I still have time, I will explore more into the neural networks method. For example, I will explore how time efficiency and result accuracy are related to the number of layers and the number of neurons in each hidden layer. With the cross validation or K-fold testing technique employed,

This project will be a trial to apply both type well construction strategies and machine learning principles, to our traditional oil and gas production data analysis. We will further apply probabilistic theory to classify the undrilled wells.

**Resources**

The data collection is going to be the production data set that is available from the website DrillingInfo. I also need to learn the mechanisms of neural networks and I tend to use Java to build a neural network for this project.

**Progress Schedule**

Oct 1 – Oct 15, data preparation, cleaning, and preprocess

Oct 16 – Oct 31, type well construction

Nov 1 – Nov 15, apply neural network to train the type well

Nov 16 – Nov 30, results improvement and other machine learning algorithms application

Dec 1 – Dec 13, project report and presentation