One Dimensional Convolutional Neural Networks for Predicting Crude Oil Properties from FTIR Spectral Data

Souvik Ta, Shahla Alizadeh, Dr. Ajay Ray, Dr. Lakshminarayanan Samavedham

Western Engineering

Chemical and Biochemical Engineering

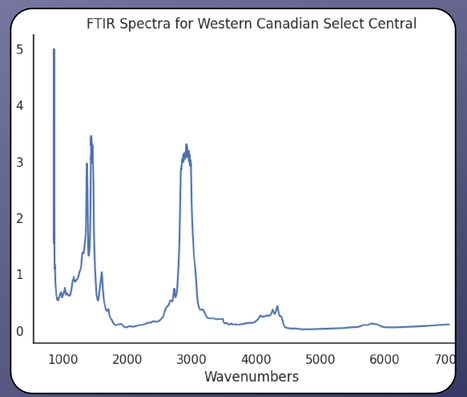
Motivation and Background

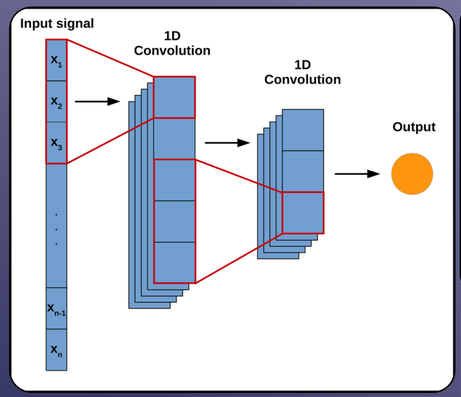
Determination of physicochemical properties of crude oil is a critical, time consuming and lab intensive process.

Developing a predictive model for the properties of crude oil which is real-time, accurate and repeatable.

FTIR Spectra has been previously used to find estimates of properties of crude oil but most of them were limited in range to specific regions or types of crude and were empirical relations

We intend to create a framework which analyses spectral data and returns accurate predictions of the properties and can be repeatable for other spectral datasets not limited to petrochemical applications





Objectives:

- To develop a real-time ,accurate and repeatable prediction model for properties of crude oil from its Spectra
- To create a general framework for spectral datasets which can be used to create predictive models

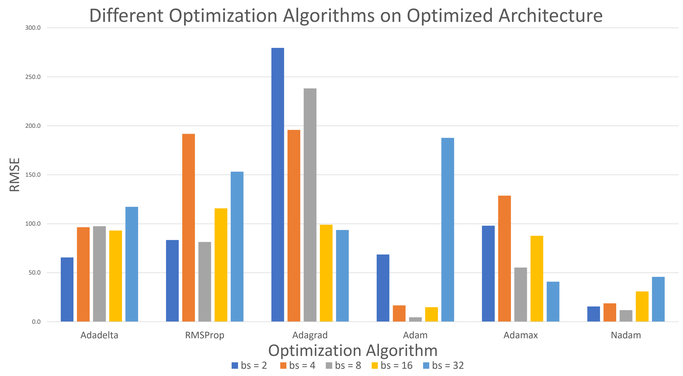
Secondary Objectives:

- Explore how the 1D CNN does feature extraction
- Gather insights into the black box nature of the 1D CNN by analysing intermediate convolutions

Dataset

The crude oil samples used in this study were obtained from seven different Canadian oil fields using a Thermo Fisher Scientific FTIR microscope

Overall, 107 samples of crude with 6366 wavenumbers each are used as attributes.



Convolutional Neural Network

- Widely used class of Deep Neural Networks that are primarily used in computer vision applications and mainly deal with images.
- Recent studies have started to use a modified CNN called a one-dimensional CNN to predict properties from spectral information. This views the spectra as a one-dimensional image
- It uses filters which convolve over the entire input map and generate an activation map using dot products. This activation map amplifies all the important features in the spectra and thus causes feature extraction
- In our application we decide to use a shallow neural network (2-5 layers) since literature tells us that deep NN's cause overfitting and higher loss in spectral applications

Hyperparameter Search

Parameters that are adjusted during the training of the algorithm manually to control the training

- Batch-size: 2, 4, 8, 16, 32
- Learning Rate: 0.1, 0.01 , 0.001
- No of Training Epochs : 100 - 300
- Optimization Algorithms: Adam, AdaGrad, Adadelta, Nadam, Adamax, RMSProp

Neural Architecture Search

Neural Architecture Search (NAS) is a technique for finding the optimal NN architecture for a given task .

It has three dimensions:

- Search Space
- Search Strategy: RandomSearch, Bayesian Optimization, HYPERBAND
- Performance Estimation Strategy

NAS can find fine tuned architectures for the specific application as opposed to using a general NN architecture and has recently achieved state of the art performance on many datasets

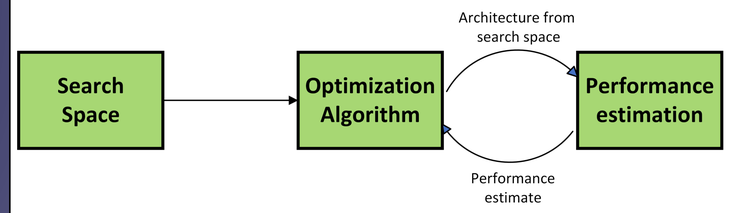
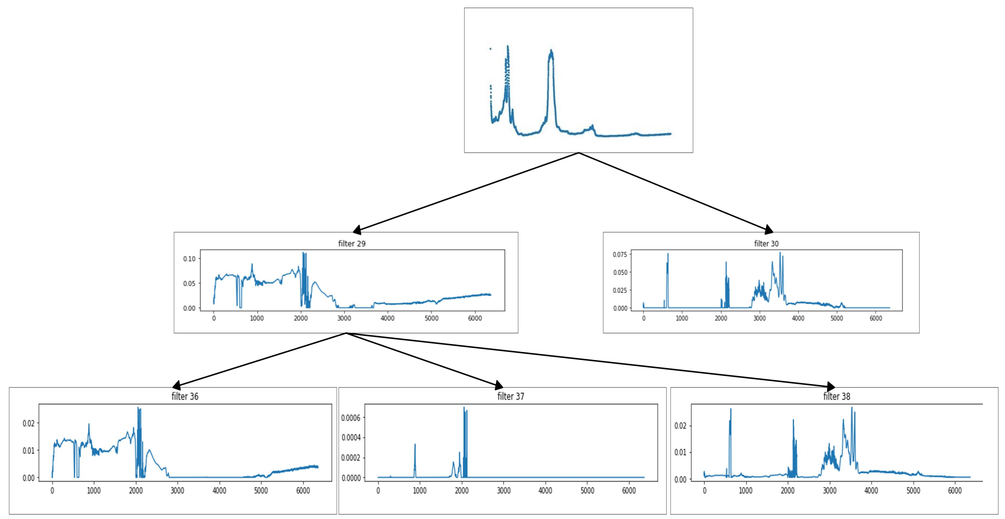


Figure: Neural Architecture Search Algorithm

Results

- For our initial hyperparameters a batchsize of 8 and a lr of 0.001 was selected as the best and the algorithm used was Nadam.
- Initially the search space was limited to two convolutional layers and the best results were for layer 1 : 32 filters and kernel size 16, for layer 2: 256 filters and kernel size 4.
- The effect of Noise injection and Dropout was also assessed and the best results were for a gaussian noise injection with std = 0.05 and a dropout of 0.4
- A pooling layer was added after each convolutional layer which decreased the computational load and allowed more architectures to be evaluated
- Adding a Pooling layer and using a HYPERBAND algorithm allowed us to navigate a larger search space and even with a upper bound of 5 conv layers the algorithm concluded that 2 or 3 layer architectures have better performance

Visualizing Convolutions



Conclusions

- In this study we succesfully demonstrate that 1 dimensional CNN's is a viable option for getting predictions from spectral data
- A general framework for creating a model to get the predictions is established
- For spectral applications NN's don't need to be deep and coupled with the fact that the network is one dimensional , the computational load is quite low and can be used for real time application
- We have investigated the benefits of different techniques like Noise Injection, Dropout and Pooling and how they positively affect the training process

Optimized Architecture