Rohan Khaitan

Curriculum Vitae

Malda 732140,
West Bengal, India
(+91) 9932562671
rohank@cmi.ac.in
Homepage
Github

Overview

Date of Birth	*April 12th, 1997*
Sex	*Male*
Language	*English, Hindi, Bengali*
Permanent Address	*Tulsihata, District: Malda, West Bengal, India*, Pin -732140
Email	*rohank@cmi.ac.in / rohankhaitan0@gmail.com*

Education

2018 – 2020 **Masters in Data Science**
Chennai Mathematical Institute *Chennai, Tamil Nadu*
CGPA: 9.72 (Recipient of Gold Medal)

2015 – 2018 **Bachelors in Statistics**
St. Xavier's College (Autonomous) *Kolkata, West Bengal*
CGPA: 8.1

2013 – 2015 **Higher Secondary Education** (XII Standard)
Ramakrishna Mission *Malda, West Bengal*
Percentage: 96% (11 th in State Board)

2007 – 2013 **Secondary Education** (X Standard)
Ramakrishna Mission *Malda, West Bengal*
Percentage: 94.7%

Industry Experience

May – July 2019 **Intellect Design Arena**, Chennai
Supervisor : Deepak Dastrala & Neha Boob

- **Topic** : **Named Entity Recognition**
- The aim was to build a model to identify the named entities for the given documents. I did all my research and finally used a Bidirectional LSTM-CNN model which worked very well. Along with this the Spacy model was also tried out.

Relevant Projects

August 2020 **Neural Style Transfer**

- Style Transfer is the process of combining the "Style" of one image with the "Content" of the other image to create an artistic imagery. The project helped me to understand how different layers behave in a Convolution Neural Network. Implemented the task using pre-trained VGG19 model in Pytorch.

April 2020 **Image Classification Using Transfer Learning**

- Idea was to use a pre-trained model and Fine Tune it to achieve a good accuracy for predicting 15 scene categories. Fine Tuning is a very crucial step in Transfer Learning. InceptionResNetV2 model after Fine Tuning worked extremely well.

April 2020 **Building a RNN from Scratch for Sentiment Analysis**

- Built a Recurrent Neural Network from scratch to identify Sentiments. The idea was to understand how the network works, Backpropagation Through Time(BPTT) and carry out a sentiment anlysis without using built-in libraries like Keras. It was a simple RNN model which helped to understand how RNN works exactly and how it differs from ANN.

Nov 2019 **Perturbed Gradient Descent to avoid Saddle points**
Supervisor : Dr. K. V. Subrahmanyam

- The goal of this Theoretical Project was to understand how PGD and Projected PGD help to escape saddle points in case of unconstrained and constrained Optimization.

Feb – March 2019 **Predicting Famine in Countries.**

- Forecasted the production, population and GDP for year 2025 from existing data for different countries.Created an index based on these three and used it to classify (XgBoost Model) whether a country will experience Famine or not.

March – April 2019 **High Frequency Time Series Analysis.**

- Worked on hourly Energy Consumption Data. Purpose was to understand different models and find out the best one. Among Seasonal ARIMA, Fbprophet and LSTM models, LSTM worked better.

September 2019 **Gibberish Text Classification**

- Understanding the context based on the words is not a good idea in this case. Here the aim was to focus on the k-grams or the characters. Basic ML models were tried out on the k-grams but CNN model based on characters was the best choice.

March – April 2019 **Dimensionality Reduction using Principal Component Analysis.**

- Some benefits of dimensionality reduction include less computation/training time, removing multicollinearity by eliminating redundant features and also in visualization of multidimensional Data. The aim was to understand the mathematics behind PCA and also to see the applications in Machine Learning.

March 2018 **Real life implementation of Logistic Regression and dealing with Multi-collinearity** (Undergraduate Dissertation)
Supervisor : Prof. Surupa Chakraborty

- The aim was to see how logistic regression helps in classification and implement it. Multi-collinearity is an important factor which effects the result significantly. The challenge was to detect it and adapt techniques to get rid of it.

Some other Projects/Relevant Assignments My Projects in Github

Project Topics

Some of the relevant topics on which I have worked on are in the following -

· **Object Detection** · **Named Entity Recognition** · **Transfer Learning**
· **Image Classification** · **Sentiment Analysis** · **Gibberish Text Classification**
· **Language Modeling** · **Word Embedding** · **Perturbed Gradient Descent**
· **Bayesian Data Analysis** · **Bandit Algorithms** · **Movie Recommendation**
· **Principal Component Analysis** · **Feature Detection & Description**
· **Regression** · **Unsupervised Learning** · **Time Series Analysis**

Relevant Coursework

Statistics	**Descriptive Statistics, Probability Theory, Statistical Inference, Non-parametric Statistical Inference, Multivariate Analysis, Linear Statistical Models, Sampling theory, Statistical Quality Control, Econometrics, Time Series Analysis, Bayesian Data Analysis**
Mathematics	**Convex Optimization, Real analysis, Numerical Linear algebra, Linear Programming Problem, Algebra I-II, Calculus (Single and multivariate), Discrete Mathematics.**
Data Science & Computer Science	**Data mining & Machine Learning, Advanced Machine Learning, Applied Machine Learning, Computer Vision, Natural Language Processing, Reinforcement Learning, Big Data with Hadoop, Regression & Classification, Information retrieval, Design & Analysis of Algorithms, Programming with C & C++, Programming & Data Structures with Python**

M.Sc Course Details B.Sc. Course Details

Technical Skills

OS	**Linux, Windows**
Programming	**Python, R, C++**
Others	**SQL**, **Hadoop**, **AWS**, **Google Colaboratory**, **HTML/CSS**, **LaTeX**, **MINITAB**

References

Dr. Madhavan Mukund
Dean of Studies
Chennai Mathematical Institute
madhavan@cmi.ac.in

Dr. Sourish Das
Associate professor
Chennai Mathematical Institute
sourish@cmi.ac.in