UTKARSH AGRAWAL | 13MA20044

MATHEMATICS & COMPUTING (M.Sc. 5Y)

EDUCATION

Year	Degree/Exam	Institute	CGPA/Marks
2018	M.SC(5YR)	IIT Kharagpur	7.27 / 10
2013	AISSCE	CBSE	86.2%
2011	SLC Exam	PABSON	89.8%

INTERNSHIPS

Affine Analytics Pvt. Ltd.   Summer Intern   MAY - JUNE 2017

Impact **:** Developed an LSTM based classification model which was integrated into Affine's text analytics product
• Implemented **paper** on "Tweet Modeling with **LSTM** Recurrent Neural Networks for Hashtag Recommendation" using Keras in Python
• Developed a sentiment analysis **application** by training a word2vec model using gensim**;** improved accuracy by **15%**
• Researched unsupervised topic modelling techniques like **Latent Dirichlet Allocation** to identify underlying topic in tweets
• Analyzed the tradeoff in accuracy with respect to modifications in preprocessing and created visualizations using bokeh library

Roofpik Pvt. Ltd.   Machine Learning Intern   MAY - JUNE 2016

Impact **:** Designed the architecture of an administrator-based chatbot and programmed the same in python using flask and firebase
• Trained a text **feature extraction model** on **wit.ai**, a text processing tool by facebook and processed the features on firebase
• Formulated a machine learning problem and collaborated with **marketing team** for effective data collection
• Distinguished sensible user queries from insensible ones using **SVM-based classification** model with linear kernel
• Hypertuned the SVM model using GridSearchCV in Sklearn for classification problems and boosted its accuracy by **5%**

PROJECTS

Online Buzz Prediction on Twitter

• Predicted mean number of active discussions; implemented a **multiple regression** model using Forward Elimination method
• Carried out multi-collinearity diagnostics using correlation matrix and VIF values; inferred 3 highly correlated group of features
• Applied **Principal Component Analysis** to extract uncorrelated features; **6** out of **11** features explained **98%** of variance
• Analyzed residuals of test dataset using standardized residual plot; **confirmed validation** of the model using normal probability plot

Application development

• Built a webcam based **motion detector** application in python; recorded and plotted its motion timings using bokeh charts
• Developed a desktop database application for scraping real estate data using Tkinter, SQLite, and BeautifulSoup in python
• Created **webmaps** with point marker feature and popup windows on **GeoJson** data using Folium and Flask
• Developed an application for a **Video Rental Store** which allows customers to create their profile, browse the CDs, inventory, check their availability, price, etc. and complete the transactions; implemented this idea in **Java** using Netbeans

Identification of cross selling opportunities

• Determined the cross-selling potential customers for an e-commerce site by **indirect segmentation** process
• Introduced multiple new features and treated missing values and outliers using mean imputation method
• Applied **k-means** clustering for uncovering the best segment of customers**;** identified number of clusters using **elbow method**

Theft analysis in electrical power corporation

• Identified defaulters responsible for stealing electricity from a power corporation using **transductive-SVM** model with radial basis kernel
• Reduced **137** dimensions to **108** by applying Principal Component Analysis and retained **98.6%** of variability
• Applied and tested other semi supervised learning techniques like transductive random forests

AWARDS AND ACHIEVEMENTS

• Won **star performer** award after the completion of my internship at **Roofpik** in June 2016
• Recipient of the prestigious **INSPIRE** Scholarship awarded by the Ministry of Science and Technology, Govt. of India
• Secured All India Rank(AIR) **2519** (Percentile - **98.3**) in JEE Advanced 2013

COURSEWORK INFORMATION

• Machine learning by Prof. Andrew N.G. (Coursera) | Introduction to Deep learning from Udacity
• Statistical Inference (Coursera) | Regression and Time Series Modelling | Statistics in R
• Probability and Statistics | Computational Statistics | Computational Linear Algebra
• Introduction to Algorithms (MIT Opencourseware 6.006) | Programming and Data Structures (Theory and LAB)
• Object Oriented System Design (Theory and LAB) | Operating Systems (Theory and LAB) | Database Management Systems

SKILLS AND EXPERTISE

Programming languages : Python, C, C++, R, Java   **Softwares :** Git, Audacity, Solidworks, Adobe Photoshop
Web Technologies : HTML, CSS, Flask   **Databases :** MySQL, PostGreSQL, Firebase

EXTRA CURRICULAR ACTIVITIES

• Member of Kharagpur Data Analytics Group; responsible for solving machine learning problem statements and mentoring juniors
• Developed a desktop application for storing the sale records of my local business in Nepal
• Member of the bronze winning hindi dramatics team at Patel Hall of Residence, IIT kharagpur
• Volunteered for Blood Donation Camp and Cloth Distribution Drive on Patel Jayanti at Patel Hall of Residence, IIT Kharagpur
• Member of cricket, table tennis,hindi dramatics and data analytics team at IIT Kharagpur
• Regular training and participation in vocal competitions