Himanshu Taneja

(979) 985 8727, 027himanshu@gmail.com
401 Stasney St Apt 304, College Station, Texas 77840

OBJECTIVE

Seeking summer internship opportunities in the fields of data science, machine learning, and natural language processing

EDUCATION

Texas A&M University, College Station, Texas May 2018
Masters of Science in Electrical Engineering, GPA: 4.0
USICT, GGS Indraprastha University, Delhi, India May 2016
Bachelor of Technology in Electronics and Communications Engineering, GPA: 74.85/100 (3rd in class)

SKILLS

Programming Languages
Most experienced with C++, Python, R, Matlab
Some experience in C, Bash, Java, Awk, SQL
Tools and Libraries
Git, Regular Expressions, Apache Spark, Microsoft Excel, NLTK, scikit-learn

ACADEMIC PROJECTS

Importance of Text in News Stories (Thesis) Nov 2016 – Present
Adviser: Dr. Ruihong Huang

- Gathering and analyzing news stories from digital platforms such as Reddit, CNN & FoxNews
- Identifying the attributes that differentiate front page news from other stories; to design a model that can suggest top news stories to the editors

Classification of Stacking Fault Energy of Alloys Sep – Nov 2016
Language used: R

- Analyzed the effect of chemical composition of Steel alloys on their Stacking Fault Energy (SFE)
- Identified most significant elements affecting SFE of alloys using T-test and Principal Component Analysis
- Trained and benchmarked the classification algorithms (Linear Discriminant Analysis, K-Nearest Neighbors, and Support Vector Machines) on the dataset

Automatic Text Classification and Summarization
Language used: C++

- Developed a text classifier using machine learning algorithms (Naive Bayes, K-Nearest Neighbors, and Support Vector Machines)
- Implemented a method to generate summary of a text article using a sentence ranking system: ‘term frequency-inverse sentence frequency’
- Techniques used: stop-words removal, stemming, term frequency-inverse document frequency, additive smoothing

OTHER PROJECTS

Image Enigma: Encrypt Digital Images
Implemented the Enigma Machine (a polyalphabetic cipher) in Python and engineered it to encrypt digital images
Person of Interest
Analyzed Enron dataset in Python using machine learning algorithms to identify persons of interest in the Enron Scandal
VocabList: A cross-platform application
Developed a cross-platform GUI application using Kivy Framework in Python to maintain a database of words for improving vocabulary
Emails in Context Menu
Designed a context menu for Openbox Desktop Environment to display emails; the application can run as a daemon to periodically get new emails

ACTIVITIES

TechSpace (Technical Club at USICT) Jan 2014 – May 2016

- Organized InfoXpression (Annual technical fest) and monthly LAN Gaming Contests
- Presented seminar on “Python Programming Language in Data Science”