Rajat Handa

Email: - handagmu@gmail.com **Work Phone: - (571)-992-9502**
LinkedIn: - https://www.linkedin.com/in/rajathandaus/ **GitHub: https://github.com/RajatHanda**

Work Experience

- **Data Scientist** *November 2018 – Present*
  Nowigence, New York, USA
  - Build duplication and relevancy classifier pipeline to find duplicate and relevant news.
  - Identify entities like person name, geo-location, organization etc. (NER) in client specific news.
  - Implement topic modelling based extractive summarizer to generate summaries about each news.
  - Extract morphological patterns from news to generate automatic alerts, news digest and reports.
  - Design REST API using Flask for connecting React.js with Python.
  - Retrieve raw news from third party terminals into MongoDB and Elasticsearch knowledge base.
  - **Tools:** Python, Flask, Jinja, Elastic, MongoDB, Digital Ocean, Mongo Atlas

- **Natural Language Processing Intern** *October 2018 – Nov 2018*
  Nowigence, New York, USA
  - Developed multi-document topic identification framework for identifying topics being discussed about each client in news.
  - **Tools:** Python, Elastic, AWS

- **Data & Operation Research Scientist Intern** *January 2018 – May 2018*
  Principal Global Equities, USA
  - Performed planning and forecasting process to rank portfolios using Universal Selection strategy.
  - Researched and evaluated various feature engineering techniques to compare stocks among different sectors.
  - **Tools:** Python, R, Tableau, Facets

- **Graduate Research Assistant** *September 2016 – December 2017*
  Engineering Education & Cyber-Learning Lab, George Mason University, Virginia, USA
  - Developed a machine learning framework for gender inference (individual (female, male), organization) on Twitter (#ilooklikeanengineer) using tweet text (LIWC) and image (CNNs) with machine learning algorithms (SVM, Random Forest).
  - Classified topics discussed across Twitter using LDA.
  - **Tools:** Python, R, SQL, D3.js, HTML/CSS, Gephi, Google Fusion Tables, AWS
  - **Publications:** https://scholar.google.com/citations?user=ZeCtXe0AAAAJ&hl=en&oi=sra

Education

- **Master's (M.S.) Data Analytics | George Mason University | (GPA: - 3.9)** *August 2016- May 2018*
  - **Courses:** Machine Learning, Natural Language Processing, Statistics, Social Media Analytics

- **Bachelor's (B.S.) Computer Science | University of Pune | (GPA: -3.55)** *August 2011- May 2015*
  - **Courses:** Data Structure and Algorithms, Database Management Systems, Advanced Database

Technical Skills

- **Languages:** Python, R, Shell Script, D3.js, HTML/CSS
- **Python Data Stack:** Scikit, NLTK, Spacy, Gensim, Numpy, Pandas, Pytorch, Keras, fastai, Mongoengine, Flask, Jinja
- **Databases** SQL, Elastic Search, MongoDB
- **Cloud Services:** AWS, Digital Ocean, Mongo Atlas

Projects

- **Problem: Classifying Toxic Comments i.e. comments that are rude, disrespectful or otherwise likely to make someone leave a discussion to help improve online conversation:**
  - Vectorized words using Word2vec and Glove.
  - Implemented multi-layer neural network using LSTM and GRU.
  - **Tools:** Python, AWS, Keras

- **Named Entity Recognizer App**: https://coolcaught.herokuapp.com