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
  - Train relevancy and categorization classifier using word embeddings with LSTM/GRU.
  - Build near-duplicate detection pipeline to filter out duplicate news.
  - Identify entities like person name, geo-location, organization etc. (NER) in client specific news.
  - Design REST API functions using Flask for making REACT apps communicate with Python.
  - Retrieve raw news from provider's REST API into MongoDB and Elasticsearch knowledge base.
  - Extract morphological patterns from news to generate automatic alerts, news digest and reports.
  - **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.
  - Implemented extractive summarizer to generate summaries for news articles.
  - **Tools:** Python, Elastic, AWS

- **Data & Operation Research Scientist Intern** *January 2018 – May 2018*
  Principal Global Equities, USA
  - Developed planning and forecasting model to forecast excess returns based on Sector-Region Neutral Strategy.
  - Researched and evaluated various feature engineering techniques to find optimal set of signals.
  - **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.
  - T**ools:** 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)** *Aug 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)** *Aug 2011- May 2015*
  - **Courses:** Data Structure and Algorithms, Database Management, Advanced Databases

Technical Skills

- **Languages:** Python, R, D3.js, HTML/CSS
- **Python Data Stack:** Scikit, NLTK, Spacy, Gensim, Numpy, Pandas, TensorFlow, Pytorch, Keras, fastai, PyMongo, Flask, Jinja, Elasticsearch
- **Data Stores:** SQL, Elastic Search, MongoDB
- **Cloud Services:** AWS, Digital Ocean, Mongo Atlas, Mongo Compass, GitHub, Bitbucket, Heroku

W - Projects

- **Toxic Comment Classifier (Kosherortoxic):**
  - This application classifies whether text entered is toxic and if it's what sort of toxicity it has.
  - Model deployed uses pre-trained Word2Vec and Bi-directional GRU having accuracy of 92%.
  - Application Link: https://kosherortoxic.herokuapp.com

- **Named Entity Recognizer (Coolcaught)**:
  - This application recognizes type of entity in text such as person name, geo-location etc.
  - Model being used is Conditional Random Field (CRF) with context-based features i.e. POS tags.
  - Application Link: https://coolcaught.herokuapp.com