

## SHYAM SUBRAMANIAN

• (508) 410-8933 • ssubramanian2@wpi.edu • www.linkedin.com/in/shyamsubramanian • github.com/shyamsubramanian

### EDUCATION

**Worcester Polytechnic Institute (WPI)**

Master of Science in Data Science

Worcester, MA

May '20

**Anna University**

Bachelor of Engineering in Computer Science and Engineering

Chennai, India

May '16

**RELEVANT COURSES:** Deep Learning, Information Retrieval, Big Data Management, Knowledge Discovery and Data mining, Statistical Methods for Data Science, Artificial Intelligence, Data Mining and Business Applications

### TECHNICAL SKILLS

|  |   |  |
|--|---|--|
| <b>Languages</b>                             | : | Python, R, C, C++, Java, JavaScript, C#  |
| <b>Machine Learning Frameworks/Libraries</b> | : | Scikit-learn, SciPy, TensorFlow, Keras, Dask, PySpark, NLTK, Matplotlib, Seaborn, Plotly, Scrapy |
| <b>Database Systems</b>                      | : | MS SQL Server 2012, MySQL, PostgreSQL  |
| <b>Big Data</b>                              | : | Hadoop, Spark, Pig, MongoDB  |
| <b>Other Analytics Tools</b>                 | : | Weka, Tableau, Google Analytics, SAS   |

### ACADEMIC PROJECTS

**Research Paper Summarization**

Spring '19

- Summarized research papers using Used TextRank and LSA methods for extractive summarization and fed the summary to Pointer Generator Network for abstractive summarization on WikiHow dataset
- Used Transfer learning for abstractive summarization from WikiHow dataset to SciSumm research paper dataset due to lack of summary dataset for research papers

**Pose Estimation using Attention**

Spring '19

- Estimated human body pose using ResNet with MAP score of 69.2 on COCO Keypoints dataset
- Explored various visual attention mechanisms including Spatial transformers and Conditional random fields

**Quora Insincere Questions Classification, Worcester Polytechnic Institute**

Fall '18

- Classified questions from Quora as Sincere or Insincere on Quora Insincere Questions dataset in Kaggle
- Converted text to vector space model using Latent Semantic Analysis, Word2Vec and Glove Word-Embeddings
- Extracted various features such as POS, sentiment analysis, profanity count and readability scores of questions
- Developed text classification system to classify questions from Quora as Sincere or Insincere using ensemble of XGBoost and RNN-LSTM algorithms with best cross validated F-score of 0.71 in Python

**PUBG Win Percentage Prediction, Worcester Polytechnic Institute**

Fall '18

- Player Unknown's Battle Ground is a famous video game. Dataset is from Kaggle with features extracted from the gameplay of gamers. Extracted statistical and ranking features based on aggregation of team players
- Developed a regression analysis system using ensembles of Random Forest and LGBM algorithms to predict the probability of winning with best cross validated MAE score of 0.026 in Python

### WORK EXPERIENCE

**Data Science Intern, Fidelity Investments, Boston, MA**

Jun '19 – Present

- Created a framework in PyTorch for Time series analysis with abstracted GPU access using pytorch-operator in Kubeflow, for scalable deployment of Machine Learning workflow on Kubernetes
- Building various models for Time-Series Multi-variate Quantile Regression analysis, including DeepAR research by Amazon, for consumer demand forecasting

**Web Developer, Worcester Polytechnic Institute, Worcester, MA**

Jan '19 - May '19

- Improved WPI website performance and accessibility using insights from Google Lighthouse and SiteImprove
- Enhanced SQL queries for faster selection and updates in Drupal PHP modules

**Member Technical Staff, ZOHO Corporation, Chennai, India**

Jan '16 - Jun '18

- Developed and maintained remote control web application, which lets IT technicians view and control any computer remotely. Technologies used are C++, HTML-5/Javascript and Java Servlets
- Developed peer-to-peer communication to speedup communication and reduce data traffic to server

### CERTIFICATIONS

- Data Science Associate from EMC Education Services
- Microsoft Certified Professional in Querying and Administering MS SQL Server 2012 Databases