ZYCUS

MANISH PRASAD

Mobile: +91-9113504749
Bangalore-560100, India
Email: manish54.thapliyal@gmail.com
GitHub: https://git.io/fhdgc

Summary

Data scientist and Full Stack developer with four years of tech industry experience in a wide range of functions Including predictive modeling, content discovery, NLP, data warehousing, and product analytics.

Education

Vellore Institute of Technology (VIT), Vellore, Tamil Nadu
M.Tech (Information Technology), CGPA - 9/10.0

University of Delhi (DU), New Delhi
BSc. (Computer Science), CGPA – 7.5/10.0

Work experience

Wipro Limited (HOLMES) *(December, 2014 – Present)*
Senior Software Engineer (Data Science)

- **Service Desk Automation:** Worked on POC for categorizing service desk tickets and predicting resolutions for a client using Machine Learning and NLP. K-Fold Cross Validation is used to select best classification algorithm among SVM, Logistic Regression, Naive Bayes etc.

- **Generic Ticket classification:** Worked on building generic ticket classifier REST api's which can select best model for predicting ticket categories and resolutions based on Grid-Search and Hyperparameter Tuning. Also building Deep Learning Classifier REST api's using TextCNN and RCNN for text categorization.

- **Explainable AI for text categorization**: Working on building different api's which can judge quality of data and explain the results of classifier (Including reasons for False Positives) for text classification.

- **Named entity recognition for chat bot:** Worked on building a multi-lingual entity extractor with different algorithms neural networks RCNN, Lstm-CRF and ensemble algorithms like Random Forest, Gradient boosting trees.

- **Cognitive Image Processing:** Worked on building a Image classifier for Table extraction and Table Detection using FCNN's (fully connected CNN's) built on tensorflow.

- **Hierarchical Text Classification:** Worked on building a hierarchical text classification that predict the hierarchy at each level of the classification and return the confidence score for each level.

- **Synthetic Data Generation:** Worked on synthetic data generation that is data generated programmatically.

  So, it is not collected by any real-life survey or experiment**.** Its purpose, is to be flexible and rich enough to be increase the variation of data in classification and named entity recognition.

***Framework & packages Used**:* Scikit-learn, Tensorflow, numpy, pandas, sklearn, matplotlib, LIME, spacy, NLTK

Technical Skills

Concepts: Machine Learning, Natural Language Processing, Deep learning, Data Analysis, semi-supervised Learning, Autoencoders, Recurrent neural networks, Convolutional neural networks, Generative adversarial networks.

Programming Languages: Python, R, Java, JavaScript, Angular, NodeJS

Databases: PostgreSQL, MongoDB, ArangoDB, MySQL, Oracles

Big Data Frameworks: Hadoop streaming, Spark

Other Technologies: HTML, CSS, YAML, Bootstrap, JQuery

Machine Learning: linear regression, hierarchical linear model, logistic regression, random forest, boosted decision trees, naive Bayes, SVM, k-means clustering, Gaussian mixture model, SVD/PCA, tf-idf, LDA, word2vec

Statistics: hypothesis testing, Bayesian inference

Academic Projects

Recommender System Using Machine Learning ***Master's Thesis- VIT***
(Guided by: Prof. Nagesh Babu)
The project aims at developing a web based application that would provide academic module recommendations on the basis of student's expertise level in different domains. The application aims to base its decisions on reliable, authentic and filtered source of information, thus saving its users from the problem of unnecessary information overload.

Toxic Comments Classification: solving the toxic comment classification challenge on kaggle. The topic is to build a multi-headed model that is capable of detecting different types of toxicity like threats, obscenity, and insult.

Exchange of classification rules in intelligent distributed systems: Classification rules are generated and exchanged among multiple intelligent agents using Matlab

Positions of Responsibility

- Managed Data Science Interns.
- Guided team of juniors.
- Handling of different analytical application
- Blogging on Data Science and its applications.

Declaration: The information furnished above is correct to the best of my knowledge

Manish Prasad *February 18, 2019*