MANISH PRASAD

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

Education

Vellore Institute of Technology (VIT), Vellore, Tamil Nadu
M.Tech (Information Technology), *CGPA - 9.45/10.0*

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

Profile Summary

Data scientist and Machine Learning Engineer with 5+ years of tech industry experience in a wide range of functions Including predictive modeling, Natural Language Processing, Computer Vision, Time series analysis and product analytics.

Work experience

ZYCUS *(May, 2019 - Present)*
Lead Data Scientist

- **Smart Desk:** Vendor query auto reply bot. Based on text classification and intent classification of the email. Based email subject, body or the attachment attached. Main functions of smart desk Query email or add invoice email.
- **Invoice Extraction**: Invoice extraction, of Headers fields, lines items and taxes and charges extraction bots.
- **Fraud email:** Detecting of potential fraud emails, when the sender of the email is not the regular contact or has a different reply to the address or When the vendor request for the change of bank details.
- **Duplicate Invoices detection:** Historical learning based model, which checks all the past invoices to see if the new one is a duplicate.
- **Invoice to Purchase order mapping:** Text similarity based bot. which automatically maps extracted line items in the invoice to the purchase order.
- **Vendor Recommended systems:** Vendor and product recommendation based on content based collaborative based filtering.

Wipro Limited (HOLMES) *(December, 2014 – April, 2019)*
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.

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. OpenCV.

Programming Languages: Python, R, C++, Java

Tools & Technologies*:* Scikit Learn, spacy, NLTK, Mysql, Tensorflow, Keras, MonogoDB, ArangoDb, NoSQL

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.

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

Manish Prasad