PRANAB KUMAR DAS

E-Mail: pranabdas457@gmail.com | LinkedIn: https://www.linkedin.com/in/pranabdas457
Github: https://github.com/pranabdas457 | +91 – 933 460 2002
Current Organization – Infosys Limited | **Current Location** - Hyderabad

Professional Summary

A dynamic Data Scientist with 3.6 years of industrial experience in Machine Learning, NLP and Software development. Lifelong learner, passionate to explore the world of data science. Holds expertise in Python, Data Analysis and Chatbot. Implemented Python frameworks to build ML models for clients, involving Business understanding, Data analysis and Text analytics.

EDUCATION

BCA: West Bengal University of Technology – DGPA 8.48/ 10.0 (**2013 – 2016**)

DIPLOMA: IGNOU – POST GRADUATION DIPLOMA IN APPLIED STATISTICS (**JAN 2020 – Present**)

SKILLS

Language Python 3, C#

Database SQL Server 2014, Oracle, Cosmos DB

Technology Data Science, Machine Learning, Scikit-Learn, Numpy, NLP, Git, Classification, Regression, Clustering, Ridge Regression, Lasso Regression, Data Analysis, Visualization, Microsoft Bot Framework, Q&A Maker Seaborn, Matplotlib, Pandas, LUIS, Asp.net MVC, SQL Server, SSIS, JQuery, JavaScript, Azure, REST API

Tools Jupyter Notebook, Visual Studio, TFS Spyder, Azure Storage

ACHIEVEMENTS & CERTIFICATIONS

2018 **Udacity, Machine Learning**
Completed Machine Learning Nanodegree from Udacity.

2018 **Andrew NG Stanford, Machine Learning**
Completed Andrew NG's Machine Learning course from Coursera.

2019 **Natural Language Processing Cert.**
Ongoing Natural Language Processing using Python from Coursera.

PERSONAL PROJECTS

2020 **Detecting malaria using cell images**
The dataset contains 2 folders – Infected and Uninfected which contains cell images. The task is to identify whether a given cell image is an infected or uninfected.

2019 **Amazon fine food reviews**
Dataset consists of reviews of fine foods from Amazon. Used classification algorithms to predict whether the review is positive or not. Used cross validation for effectiveness of the model.

2019 **MNIST Digit Recognizer**
It is a classic dataset of handwritten images. Used this dataset to visualize the cluster by reducing dimension. Also used various algorithm for classification task.

2018 **EDA and visualization using Python**
Investigated a dataset using Python to discover patterns, spot anomalies, check co-linearity, test hypothesis and to check assumptions with the help of summary statistics and graphical representations.

EXPERIENCE [reverse chronological order]

Project Title: Chatbot development using **Microsoft Bot Framework v4**. The objective of the project is to develop a master bot and 4 skills bots (child bot) for different departments. All the child bots have their own LUIS. The master bot is the main bot which takes inputs from the user and redirects the query to the respective child bot.

- ✓ Used dispatch to create and evaluate LUIS models and to identify intent.
- ✓ Developed web jobs in Azure cloud to update database from bot.
- ✓ Implement language detection based on user input.
- ✓ Used NLP models for detecting Hate Speech.
- ✓ Implemented spell check, auto suggestions and feedback functionality.

Technologies – C#, Bot Framework V4, REST API, Cosmos DB, Azure Cloud, Git, SQL Server, GIT.

Project Title: Identifying different areas of improvement from customer reviews and feedback to provide better services.

- ✓ Analyzed customer reviews and feedback to analyze their problem and provide services accordingly.
- ✓ Gathered data from their feedback portal. Cleaned the text and remove stop words and special characters properly.
- ✓ Used topic modeling technique to find different topics.
- ✓ Used NLP techniques like word tokenization, **TF-IDF** and **WORD2VEC** models to identify important keywords.
- ✓ LDA and LSA algorithm was used to identify different topics.
- ✓ Used Word Cloud to visualize the topics.

Technologies: Python, Pandas, Scikit Learn, Matplotlib, Seaborn, Jupyter notebook.

Project Title: Developed an email classifier.

- ✓ Developed a model to classify spam and non-spam emails using Natural Language Processing concepts.
- ✓ Applied Text Analytics using the Python libraries NLTK, in cleaning the text from punctuations and special symbols and converted the text data into TF-IDF and Word2Vec model.
- ✓ Applied **Naïve Bayes** algorithm on the pre-processed data to classify as spam or non-spam.

Technologies: Python, Pandas, Scikit learn, NLTK, Matplotlib, Seaborn, Jupyter notebook.

Project Title: Developed a web application for maintaining the field technicians' crew and shift details.

- ✓ Designed the UI using HTML, CSS, and KENDO UI
- ✓ Used Ajax, JavaScript and JQuery in client side.
- ✓ Developed complex stored procedures.

Technologies: C#, Ajax, Asp.Net MVC, SQL Server, Kendo UI.