Hardik Jivani

hardik.jivani@nyu.edu | (551) 228-8614 | linkedin.com/in/hardikaj96 | github.com/hardikaj96

EDUCATION:

New York University, Master of Science in Computer Science | GPA – 3.6/4 May 2020

(Courses: Database Systems, Big Data, Machine Learning, Design & Analysis of Algorithms, Cloud Computing, Computer Vision)

Mumbai University, Bachelor of Engineering in Computer Engineering | CGPA – 8.7/10 May 2018

(Courses: Artificial Intelligence, Machine Learning, Image Processing, Soft Computing, Parallel and Distributed Systems)

TECHNICAL SKILLS:

Programming Languages: Python, SQL, PHP, Node.js, Javascript, jQuery, C, Java

Python Packages: scikit-learn, nltk, pandas, numpy, scipy, statsmodels, scrapy, matplotlib, seaborn, spacy, selenium

Cloud Technologies: Amazon AWS, EC2, Lambda, Lex, Elastic Search, DynamoDB, Kinesis, SageMaker, Cognito, Rekognition, RDS

Tools/Technologies: Django, Flask, Pyramid, Laravel, Git, Angular.Js, React, Computer Vision, Docker

Big Data/Databases: Hadoop, Kafka Streaming, PySpark, SparkSQL, SparkML, MongoDB, MySQL

EXPERIENCE:

Data Science Intern, Applied Informatics Inc, New York, NY May 19-Aug 19

- Collaborated comprehensively with data science team and software development team to ensure that solutions meet business context and pass valid hypothesis testing
- Implemented scalable and fault-tolerant well-documented and robust code across a distributed architecture using Docker for scraping clinical trials data
- Track exceptions and completion status of scraping process and send notifications using Chat API
- Performed Exploratory Data Analysis on the Clinical Trials data to gain a deeper understanding of the problem and generate data-driven insights
- Developed the next-generation automation script for performing Exploratory Data Analysis with matplotlib and seaborn
- Performed Statistical Significance test on each attribute with the target feature of any given dataset
- Implemented predictive models for result prediction of clinical trials based on the clinical trial title

Research Intern, **Tata Consultancy Services,** Mumbai, India Jun 17-Dec 17

- Implemented Entity Recognition and labelling intent using Spacy by training various sentences of a Library Management System and achieved 88% validation accuracy
- Designed Web Application using the above model in Django which predicts the user's intention and the entities based on the input sentence as per the system specified, which also adds new systems and trains the model based on user's training data.
- Collaborated with interns in Image Classification problem (tensorflow) which had across high processing time & computation

PROJECTS:

Smart Door Authentication System (AWS, Kinesis Video/Data streams, Lambda, Rekognition, DynamoDB) Oct 19-Nov 19

- Developed a system to identify the visitors using AWS Kinesis Video Stream and AWS Rekognition that authenticates visitors and provides them access to a virtual door
- Implemented Lambda Functions that sends OTP to the known visitors and sends a notification to the owner for the unknown visitors. Created multiple tables in DynamoDB to store visitors information and passcodes with TTL feature

Dining Concierge Chatbot (AWS, Lex, Lambda, Cognito, SageMaker, DynamoDB, Yelp API)
Sept 19

- Developed a web-based ChatBot to recommend restaurants to the customers, based on cuisine and location preferences
- Extracted entities from the user's input using Amazon Lex and notified user with restaurant suggestions with the help of yelp API through an email and text message

Professional League Quidditch player classification (Python, scikit-learn, pandas, scipy) Mar 19-May 19

- Implemented Missing Value Imputation; Feature Conversion, Extraction and Generation on the dataset
- Analyzed comparatively efficient Machine Learning Models such as Logistic Regression using SMOTE, Decision Tree Classifier and Multi-Layer Perceptron Classifier on Quidditch player dataset to get the training and validation accuracy
- Tuned the models with hyper-parameters and re-sampling method.
- Achieved best validation accuracy for Logistic Regression with SMOTE and later tested on the Test Dataset

PUBG Rating System & Analysis (Python, PySpark, Kafka, SparkSQL, Node.js, MongoDB) Mar 19- May19

- Simulated PUBG data with Kafka Streaming to produce each match data and SparkStreaming to consume the stream
- Developed PUBG match rating algorithm to keep rating relative with the recent activity of each player in matches and update data in the MongoDB. Built MongoDB Pipeline to display statistics on the PUBG ratings website built using Node.js

Stock Market Prediction (Python, Matplotlib, NumPy, Statistics) Jan 18-May 18

- Implemented back-propagation training algorithm with k-folds cross validation which takes different normalized factors to predict the next day's opening NIFTY100 index and visualized the correlation b/w input with opening NIFTY100 index

PUBLICATIONS:

"**Forecasting Indian Stock Market Using Artificial Neural Networks**", presented in International Conference on Computing, Communication, Control and Automation. Publisher-**IEEE** (8697724) Published: April 25, 2019