|  |  |  |
| --- | --- | --- |
| |  | | --- | | Suraj Kesavan  42/6, Subhiksha Apartment, Kannagi Street, Chennai-600094  +91-8939493808 – suraj.msec@gmail.com | |  | |

|  |  |
| --- | --- |
| **About** | Working currently as a Data Scientist, having experience in Data Analysis along with experience in web application development using BPM and Digital Marketing tools and graduate of PGDM in Data Science |
| Experience | **Technology Analyst – Tata Consultancy Services (Mar 2018 - Current)**   * Currently working on ABR(Analytics Based Repair). Telematics, Engine Claims and Complaints data were analyzed. **pySpark** on **Databricks** was used which was hosted on **Azure. Random Forest** model was built with accuracy of 86%. Recall was used as a performance metric. * Customer usage data was preprocessed and analyzed to get insights about the usage using **python** and its **pandas** library * Built a **Decision tree classifier** model to determine whether or not to trigger offer to corresponding customer * model was integrated into **pega marketing** using **sklearn2pmml** package and imported into pega marketing using the **pmml** file * Building campaigns using strategies and decision rules to trigger appropriate offers to customers via corresponding channels * Worked on building a Ticket Handling application for Telecom Domain - Worked on **services** and **listeners** to provide connection to different systems and designed most of the **database** schema for the application, also went onsite to Hungary office to setup and deploy the Production Environment for the same * Application Monitoring was setup in python, case data taken from database was analyzed and reports were created and sent to client * **Webpage Analysis** – Data from log files were collected and analyzed for gathering insights on the time spent on each of the application(Insurance domain) screens, required information was got using regex library and converted to a dataframe in **python** using **pandas** library, **matplotlib** was used for visualization. This information was used to secure additional enhancements on those application screens from the client.   **Senior Systems Engineer – Infosys (Aug 2014 – Feb 2018)**   * Building **web application** for back office for **BPM** processes in a banking domain like **Wealth Management** and **Conflict Management** using **Pega** tool, both applications were built from design phase and successfully deployed * Completed training in JAVA and reskilled to Pega |
| Projects | **Stock Price Prediction**  **Tools Used**; Python sklearn Libraries, Random Forest Regressor, Support Vector Regressor, TimeSeriesSplit, tensorflow and keras libraries, adam optimizer, activation functions, CNN, LSTM, Attention   * Predicting the stock price using machine learning and deep learning technique and analyzing which model works best * Univariate and Multivariate prediction of data for next week using available data from previous week and previous two weeks * Walk forward validation method was used to evaluate the model  Stroke Prediction (Healthcare) **Tools Used**; Python sklearn Libraries, Tree-based algorithms (Random Forest, XGBoost), Classification Metrics (Precision, Recall, Confusion Matrix), SMOTE Library, cross\_val\_score with Stratified KFold   * Predicting whether or not a person has a chance of getting a stroke on a McKinsey Stroke Prediction Dataset * Exploratory Data Analysis on factors that lead up to the stroke * Imbalanced class handling with methods like oversampling and smote sampling  Flight Ticket Price Prediction **Tools Used**; Python sklearn Libraries, Regresion-based algorithms (Random Forest, Gradient Boost), Regression Metrics (RMSE), GridSearchCV , cross\_val\_score   * Predicting price of flight ticket using regression techniques on a machine hack dataset * Data Wrangling to extract information from uncommon field formats  Clustering and Segmentation **Tools Used**; Python sklearn Libraries, Agglomorative and Heirarchical Clustering, Dendrogram, K-means Clustering   * RFM analysis on transactional data of a store * Segmentation of customers based on cluster characteristics * Marketing actions corresponding to the cluster characteristics were tailored |
|  |  |
| Education | Post Graduate Diploma in Data Science Praxis Business School  December, 2020  CGPA – 6.15/8 Graduate in Bachelor of Engineering in Electrical and Electronics Meenakshi Sundararajan engineering college – Anna University  (2010 – 2014)  CGPA – 6.73 High School Padma Seshadhri Bala Bhavan – CBSE Board  (2008 – 2010)  Marks – x: 76 %; xii: 72.8% |
| Certifications | * Deep Learning Specialization - Coursera * Pega Decision Consultant |
| Portfolio | * <https://github.com/SurajK02> * <https://www.linkedin.com/in/suraj-kesavan-5a396199/> |
|  |  |
|  |  |