Shrikant Kendre

Associate Data Scientist – Cognizant

Email: skendre@dons.usfca.edu **Phone:** +1(650) 680 9636 **Address:** San Francisco, California
GitHub: github.com/shriawesome **LinkedIn:** linkedin.com/in/shrikant-kendre-2941a6143
Portfolio: shriawesome.github.io

SUMMARY

3 years of extensive experience in **modelling**, **predictive & quantitative analysis** and **natural language processing, Deep Learning and Data engineering**. Worked for clients across **Healthcare, retail** and **IT** domain with primary expertise in developing & calibrating **regression**, **classification** and **clustering** frameworks along with **text mining** models using **Python** & **SQL**. Experienced in working across multiple stages of data analytics project starting from data acquisition/ data preparation to hypothesis building/ modelling followed by evaluation & interpretation and deployment with post production model optimisations.Also worked on creating data pipelines on cloud platform like AWS to migrate data on cloud platform. Have completed BE in CS from Maharashtra Institute of Technology, Pune.

WORK EXPERIENCE

02/2021 - 08/2021

Software Developer, PhlexGlobal Pvt. Ltd.

Developing and calibrating machine learning / Deep learning models for healthcare domain and Pharmaceutical industry. Also worked on Optical Character Recognition using Pytesseract, PDFTron and Azure OCR for converting PDFs to Text Documents later to be used for Named Entity Recognition.

09/2018 - 12/2020

Associate Projects, Cognizant Technology Solutions

Developing and calibrating machine learning / Deep learning models for different application and domains like healthcare and IT-Products. Also worked with creating python scripts for data visualisation for Banking domain along with experience in creating data pipelines on AWS platform for retail client.

EDUCATION

08/2014 – 05/2018

Bachelor of Engineering in Computer, Maharashtra Institute of Technology - 3.35/4 (G.P.A.)

Achievement **:** Grade ‘A’ (First Class with Distinction, SPPU accreditation)

TOOLS & TECHNIQUES

Tools: Python(Pandas, numpy, scikit-learn, matplotlib, seaborn), C++, Jupiter Notebook, Google colab, Scikit-Learn, TensorFlow(Keras), AWS, Azure Cloud, Git, PySpark, SQL & MongoDB.

Algorithms: Linear Regression, Logistic Regression, SVM, Decision Trees, Random Forest, Convolutional Neural Networks, Recurrent Neural Networks, NLP.

PROJECTS

Worked on PhlexEview, performing NER on Summary of Product Characteristics(SMPCs) documents
PhlexGlobal Pvt. Ltd.
Client wanted to convert physical SMPC documents into digital records that can be later used to perform quick analysis
Physical documents were converted manually to digital formats(Searchable/ Scanned PDFs), that were later used for the task of Optical Character Recognition to convert into text documents to be used for the task of NER. Corpus of important entities were extracted from the document using Deep Learning Model, which helped Pharmaceutical Industry to transform the processing of electronic Trail Documents and improve processing with enhanced oversight and quality.

Improve the Ad Click Prediction Accuracy for Search Engine(IT Product)
COGNIZANT TECHNOLOGY SOLUTIONS
Client wanted to improve the previous model used for suggesting advertisements as a search result.
Researched on various NLP algorithms for information retrieval to estimate click probability of advertisement given as a search result.The solution was implements using a CNN and significantly impacted business unit with a 10% improvement in click through rate. Deployment via TF-serving added an additional improvement in query response time by 5 ms for search engine.

Medicare Analytics using Health Outcome (Healthcare)
COGNIZANT TECHNOLOGY SOLUTIONS
Client wanted to bolster Customers Marketing team by providing them with apt population to improve their medicare plans.
Developed analytical and ML model that assist customer marketing team by targeting right and influential population to improve the star ratings of medicare plans by addressing specific measure. Various feature selection techniques were implemented and based on the result various ensemble models were trained and the result was evaluated over confusion matrix and auc.

Data Validation and Visualisation using different Macro Economic Variables (BFS)
COGNIZANT TECHNOLOGY SOLUTIONS
Client wanted to transfer data from oracle to sql and wanted to automate the task of data cleaning along with performing data visualisation on the SQL data.
Wrote a python script that automated the entire task of performing data validation across ORACLE and SQL and, data visualisation based on different MEVs and the result was automatically generated in the form of a PDF. This majorly impacted the business unit by saving lot of time and labour required previously for manually performing data validation using excel and, also the plots were used readily by business analytics team to see how a particular MEV is behaving over a period of time.

Data pipelines to shift data from local database to RedShift(Retail)
COGNIZANT TECHNOLOGY SOLUTIONS
Client wanted to migrate data from its traditional database and variegated sources to on cloud AWS Redshift service to help the analytics team for further visualisation and analysis of the data by Analytics team.
Gathered data from various services like and automate the task of Data Migration to AWS Redshift and validated the data via data visualisation using Athena and querying in Redshift. This significantly affected the business unit by fast processing and retrieval of data using AWS Services and also the sns service by AWS helped the client significantly by updating the team whenever the new data came in that resulted in quick analysis and use of data. Different services by AWS made handling and managing the data easy and solved the problems faced by on-premise local databases.

PERSONAL PROJECTS

Automatic Detection of Hateful Text in Online content using Twitter API - BE Project
Maharashtra Institute of Technology - Pune
Agenda of the project was to tackle an inexorable problem of hateful/ offensive content on online platforms that sparked many social issues and riots in public.
Built an intermediate module between user and the twitter platform that monitors tweets and prevents user from tweeting anything offensive/ hateful online, hence maintaining and monitoring the decency of the platform. The solution was implemented using multi-class Classification via Naive Bayes, SVM and Logistic Regression and evaluated over 'precision' and 'recall' along with ROC curve. The proposed solution can help many social media platforms in maintain

Santander Customer Satisfaction - BFS
KAGGLE
Santander Bank wanted to identify dissatisfied customers early in their relationships and by doing so would allow Santander to take proactive steps to improve a customers happiness before they lose them.
Developed different ML models and performed various feature selection techniques to drastically reduced feature space that in turn boosted the model performance by keeping the accuracy of the model unharmed. The results were very promising and model performed significantly well on both the train set(accuracy of 80%) and test set(accuracy of 81%) along with a proper precision and recall

Image Classifier using Deep Learning - Docker based app
CIFAR Dataset - **https://hub.docker.com/r/shriawesome/cifar_classification**
Based on CIFAR-10 dataset that contians 80 million tiny images of 10 object class, the problem statement is to create DL model that can classify a given image.
Developed DL model implementing CNN as a classifier to identify the image, working on the dataset model the solution provided a training accuracy of 73% and a validation accuracy of 74% and the result was also evaluated based on confusion matrix. This model was later built into a Flask App and was deployed using Docker.

ACCOMPLISHMENTS

Publication Automatic Detection of Hateful text in Online Content using Twitter API - BE Project(2019)
https://arxiv.org/abs/1809.08651

CERTIFICATIONS

Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimizaition - Coursera
https://www.coursera.org/account/accomplishments/certificate/EQRT282WA4PH

Neural Networks and Deep Learning - Coursera
https://www.coursera.org/account/accomplishments/certificate/SMC6SNN9WRVA

Certified C Programmer - SEED Infotech Ltd
Passed 'C programming' certification with 'A' grade