ORACLE
Certified Associate

Hrusikesh Panda
Sri Sai paradise,3rd floor, Flat no. 301,
Ashraya Layout, 1st stage, 7th A cross,
Mahadevapura, Bangalore-560048
Mobile Number : +919886036753
Email: hrusikesh.aiml@gmail.com

PROFESSIONAL EXPERIENCE

Having 9+ years of diversified IT experience in analysis, design and development, Team building and leadership experience in Data Science (Python, Statistics, Machine Learning, Deep Learning & NLP) and Data Engineering (Oracle, Data Warehousing using Informatica Powercenter, Teradata, MongoDB, PGSQL, Tableau and Big data using Spark & HIVE).

DATA SCIENCE EXPERIENCE

- Work closely with the Data science Delivery team to deploy end to end solutions for Integrated Data Model.
- Have developed effective Talent Scouting, Team Building and Organizational Leadership skills.
- Understanding Business Problem Using **Python** and **Bigdata** concepts (HDFS, MAPREDUCE, HIVE and PYSPARK) to solve business Problems and successful implementations of solutions.
- Possess expertise in Data science (**Machine Learning, Deep Learning & NLP**) using Python.
- Possess expertise in implementation of **Regression** and **classification** algorithms (Linear Regression, Multiple Liner Regression, Logistic Regression, KNN, Naive Bayes classifier, Decision Tree, Random Forest, SVM, K-Means, Hierarchical clustering, ANN, CNN, RNN, Computer vision, Natural language processing(NLP), Recommendation systems, LSTM, Time series analysis) using scikit learn and tensor flow libraries.
- Doing exploratory analysis and took inferences by visualization the data using **python** and **Tableau.**
- Improvised the quality of data by removing inconsistent data, missing values and label the data using **feature scaling** and **Ensemble Learning techniques** to optimizing the model.
- Removed insignificant variables using Dimensionality reduction techniques (**PCA**, **t-SNE**, **UMAP**).
- Good at **statistical** and **mathematical** knowledge.
- Strong knowledge of **Banking**, **Retail**, **Logistic**, **Healthcare** and **Education** domain.

DATA WAREHOUSEING EXPERIENCE

- Strong ETL Knowledge of **Data Integration**, **Master Data Management**, **Data Profiling** and **Data quality processes**.
- Experience in working with using Power center **web service**, **partitioning** and **PDO**.
- Good in writing complex SQL queries according to business rules and Knowledge in **OLAP** Function and **Materialized** views.
- Possess expertise in **Database** as well as **Data warehousing concepts**, skilled at progressing from problem statement to well **documented design**.
- Possess expertise in implementation of **SCD type1, SCD Type2, SCD type3 and CDC** and knowledge in Implementation **of SCD4 and SCD6**
- Experience in working with **complex** mappings using **Mapplets**, **expressions, XML parser transformation, routers, lookups, aggregators, filters, updatestatergy, SQL, joiners and union transformations** in **Informatica.**
- Excellent knowledge and experience in **Data Warehouse Development life cycle, Dimensional Modeling**, **Repository Management** and Administration, Implementation of **STAR**, **Snowflake Schemas** and **Slowly Changing Dimensions**.

EMPLOYMENT HISTORY

- Worked as Tech Lead in OnDemand Agility (client **Credit Suisse Bank**), Pune from **Oct-2020** to till date.
- Worked as **Lead Data Scientist** in **Focus Edumatics**, Bangalore from **10-Sep-2018** to **18-Oct-2020**.
- Worked as **Module Lead** in **Persistent Systems Limited** (Client **Wells Fargo**), Bangalore/Hyderabad from **09-Apr-2015** to **07-Sep-2018**.
- Worked as **Software Engineering Analyst** in **Accenture**, Bangalore from **03-Aug-2014** to **08-Apr-2015.**
- Worked as **Software Engineer** in **UST Global**, Bangalore/Trivandrum from **01-Sep-2013** to **01-Aug-2014.**
- Worked as **Software Engineer** in **Anthem Inc**, Bangalore/Gurgram from **02-May-2011** to **30-Aug-2013.**

EDUCATION

- **MCA** from **SATHYABAMA UNIVERSITY, CHENNAI** in 2011**.**

PROJECT DETAILS

Projects	:	**LIBOR Document digitization process**
Domain	:	**Banking**
Client	:	**Credit Suisse Bank**
Duration	:	**Oct 2020 – Till date**
Technology	:	**Python, NLP, Bigdata, impala, kudu, Unix, Qlik and Finance portal**
Role	:	**Tech Lead**

LIBOR Document digitization process: In this project documents are complex, unstructured, non-homogeneous varied formats, hard to read, hard to find specifies. Digitizing, introspecting, classification and keyword/entity extraction of /from documents. below are the steps used for document digitization process.

- NLP classification algorithm to identify clause.
- NLP Extraction algorithm for keyword extraction.
- Big data store for digitized data, supporting metadata, extracted entities and clause scoring management.
- Orchestrator to link, connect and run.
- Visualization using Qlik and finance portal. workflow using flowable (open source) and chat bot for Q&A interaction.

Responsibilities:

- Understanding the requirements from business and SME's and assigning a task in a team.
- Doing data analysis and cleansed the data.
- Implemented NLP classification algorithms for identifying clause and NLP Extraction algorithm for keyword extraction.
- Using python connect to Impala and kudu table to store the data in a structural format.
- Using Unix and python script to automate the process and involved in UAT and production deployment process.

Projects	:	**AI-Tutor**
Domain	:	**Education , Healthcare ,HR Analytics**
Duration	:	**Sept 2018 – Oct 2020**
Technology	:	**Python, Google API, Google Dialogflow, Google Cloud, Machine learning, Deep learning and NLP**
Role	:	**Lead Data scientist**

AI-Tutor is an e-learning Bot here we intend to create a Bot which automates the current teaching system. Here the AI Bot would be smart enough to teach a particular topic from the subject selected by the students. Also, when the student responds to the tutor Bot, Bot will analyze the response and guide the student appropriately as per the student's ability to grasp the content.AI-Tutor would teach the topic by slowly increasing the level of difficulty of questions and continuously engaging the student to actively participate and drive the conversations ahead.
For building the Bot we used Google's natural language understanding developer framework for building conversational experiences. DialogFlow needs to be trained on the dataset to attain a machine learning capability which understands the intent and context of what a user says in order to respond in the most useful way.

DialogFlow lets you build conversational interfaces on top of your products and services by providing a powerful natural language understanding (NLU) engine to process and understand natural language input.

Responsibilities:

- Understanding the requirements from SME's and assigning a task in a team.
- Doing data analysis and cleansed the data.
- Implemented the Chatbot using Google DialogFlow.
- Automated the whole Chatbot process using **python** code, which reduced the time from 4days to 5min for each bot.
- Implemented Decision maker algorithm using NLP text similarities algorithms.

Sentiment Analysis on AI-Tutor data: Here we are performing sentiment analysis on AI-Tutor data to identify response delay, spelling error, technical error, negative sentiment, positive sentiment from each session on daily basics to generate the report. .the main goal of sentiment analysis is to improving the quality of teaching process and reduced the manual efforts.

Responsibilities:

- Made exploratory data analysis and cleansed the data.
- Used NLTK and Vader sentiment to performing sentiment analysis.

Handwritten Text Recognition System for AI-Tutor using OCR & HTR: Here AI-Tutor would teach the topic by slowly increasing the level of difficulty of questions and continuously engaging the student to actively participate and drive the conversations ahead with using white board which having images and hand written text.

Handwritten Text Recognition (HTR) systems transcribe text contained in scanned images into digital text. We build a Neural Network (NN) which is trained on word-images from the IAM dataset. As the input layer (and therefore also all the other layers) can be kept small for word-images, NN-training is feasible on the CPU (of course, a GPU would be better). This implementation is the bare minimum that is needed for HTR using TensorFlow.

Responsibilities:

- Developed the application with the help of Pytesseract
- Conducted unit testing of the application
- Deployed into Ducker with the help of flask

POC on Topic Modeling: It becomes difficult to know the context of discussion when multiple users are chatting on a common platform. It's very tedious job to go through the discussion and manually figure out the discussion topics. Also manual method is error prone. We worked on one POC where we collected the discussion of all the new joiners of the company. We analyzed the texts to find top N optimal discussion topics. This helped the trainer/ HR team to take an informed analytical decision

Responsibilities:

- Cleaned the text data
- Used Gensim for the Topic Modeling
- Used perplexity and topic coherence to find goodness of the fit
- Presented the topic with good visualization to the HR and Training team

Projects	**:**	**Product Recommendation, Value prediction, Loan defaulter prediction, Customer Satisfaction, Customer Transaction Prediction, Customer Due Diligence ,Customer feedback analysis, Risk Analysis , Lima chatbot and Wells Fargo Robust spends Forecasting**
Client	**:**	**Wells Fargo**
Domain	**:**	**Investment Banking**
Duration	**:**	**Apr 2015 – Sept 2018**
Technology	**:**	**Python, Rasa, Machine Learning, Deep Learning & NLP**
Role	**:**	**Module Lead**

Wells Fargo & Company is an American multinational banking and financial services holding company which is headquartered in San Francisco, California, with "hubquarters" throughout the country. It is the fourth largest bank in the U.S. by assets and the largest bank by market capitalization. Wells Fargo is the second largest bank in deposits, home mortgage servicing, and debit cards.

Wells Fargo Robust spends Forecasting: Wells Fargo is looking into building onto its Procurement practice, it has been able to build a robust spend Forecasting tool Using the existing spend data, calculate spend and savings projection for future. The company is interested in knowing its Procurement space. The key metrics are:

- Spend (Drill down by region, country, category, vendor).
- Spend Projection (based on past 10 years of data).
- Savings Projection (calculate the current year's projected full spend and assume a 6.5% of savings rate)
- Add additional metrics (based on the data provided).
- Establish a model that will be able to classify areas where we could expect the spend to go up by At least 10% based on last year's Jan-May and Jun-Dec data.
- Procurement is in need of a chat Bot to be able to query it usual spend and savings data on
    - What is my overall spend for 2017?
    - How much have I spent in Commercial category in June 2017?
    - What will be the 2017 spend in R&D Category?

- What will be the 2017 projected savings for Corporate category?

Responsibilities:

- Done exploratory analysis and took inferences by visualization the data.
- Improvised the quality of data by removing inconsistent data, missing values and label the data using feature scaling.
- Removed insignificant variables using Dimensionality reduction technique.
- Performed algorithms like Prophet, LSTM, ARIMA etc.
- Deployed into Ducker with the help of flask.

Wells Fargo Lima Chatbot : Chatbot is a machine-learning based conversational dialog engine build in Python which makes it possible to generate responses based on collections of known conversations. The program selects the closest matching response by searching for the closest matching known statement that matches the input, it then returns the most likely response to that statement based on how frequently each response is issued by the people the bot communicates with.

Responsibilities:

- Developed the Chatbot from scratch.
- Trained the model for predefined set of QnA.
- Made synthetic stories from the trained QnA.
- Unit tested on specific scenarios.
- Deployed the chatbot on UAT and subsequently on Production.
- Created logfile for easy debugging the issue.

Customer Due Diligence (CDD) prediction: CDD information comprises the facts about a customer that should enable an organization to assess the extent to which the customer exposes it to a range of risks. These risks include money laundering and terrorist financing. Primary goal of CDD enables Wells Fargo to know its customer understand the nature and purpose of customer relationship to develop a customer risk profile and reasonably predict the types of transaction in which a customer is likely to engage and determine when transition are potentially suspicious. Organizations need to 'know their customers' for a number of reasons:

- to comply with the requirements of relevant legislation and regulation
- to help the firm, at the time the due diligence is carried out, to be reasonably certain that the customers are who they say they are, and that it is appropriate
- to provide them with the products or services requested
- to guard against fraud, including impersonation and identity fraud
- to help the organization to identify, during the course of a continuing relationship, what is unusual and to enable the unusual to be examined;
- if unusual events do not have a commercial or otherwise straightforward rationale they may involve money laundering, fraud, or handling criminal or terrorist property
- to enable the organization to assist law enforcement, by providing available
- Information on customers being investigated following the making of a suspicion report to the Financial Intelligence Unit (FIU).

Responsibilities:

- Doing Customer interaction and understanding the requirement.
- Getting the requirements from client and distributing the task among the team members.
- Done exploratory analysis and took inferences by visualization the data.
- Improvised the quality of data by removing inconsistent data, missing values and label the data using feature scaling.

- Removed insignificant variables using Dimensionality reduction technique.
- Used Gradient descent algorithm to minimizing the cost function.
- Handled imbalanced data using SMOTE and Performed Random Forest, XGBoost, Logistic Regression, and LightGBM to find fraud customers.
- Optimized the accuracy parameters with 95% accuracy with using Random Forest and XGBoost
- Deployed into Ducker with the help of flask

Risk Analysis:

Loan IQ Prediction: LoanIQ allows you to instantly identify high-risk loans and reduce overall default exposure. The LoanIQ and Market Risk Scores assist you in making a review decision based on the level of collateral risk associated with a loan. Whether you need to assess one loan or thousands, LoanIQ delivers the answers you need fast to make the most informed loan decisions. Primary goal of Loan IQ prediction is

- Easy-to-use and designed to work with your existing loan systems and processes
- Select the highest risk loans for quality control and due diligence
- Monitor portfolio performance to aid with loss mitigation and retention programs
- Investor protection with insurance option.
- The industry's only patent-pending collateral risk predictive model
- Loan scoring based on the widest range of predictors
- Instant identification of property over-valuation
- Filter out high-risk loans by drilling down on red-flagged indicators to assess quality of loans against your underwriting criteria.

INFOLEASE prediction: Wells Fargo providing the loans to customer for keeping the collateral as a lease. We need to identify the highest risk loans for quality control and due diligence based on collateral value.

REALM prediction: Wells Fargo providing the loans to real state customers. We need to identify Instant identification of property or collateral over-valuation.

Responsibilities:

- Done exploratory analysis and took inferences by visualization the data.
- Getting the requirements from client and distributing the task among the team members.
- Improvised the quality of data by removing inconsistent data, missing values and label the data using feature scaling.
- Removed insignificant variables using Dimensionality reduction technique.
- Performed algorithms like KNN, NAÏVE BAYES, Random forest, XGBoost, LightGBM and obtained optimal sensitivity and specificity with logistic regression.
- Optimized the accuracy parameters with 92% accuracy with using Random Forest and XGBoost.
- Deployed into Ducker with the help of flask.

Wells Fargo Product Recommendation: Ready to make a down payment on your first house? Or looking to leverage the equity in the home you have? To support needs for a range of financial decisions, wells Fargo offers a lending hand to their customers through personalized product recommendations.

Under their current system, a small number of wells Fargo customers receive many recommendations while many others rarely see any resulting in an uneven customer experience. Here, wells Fargo is challenging to predict which products their existing customers will use in the next month based on their past behavior and that of similar customers.

With a more effective recommendation system in place, wells Fargo can better meet the individual needs of all customers and ensure their satisfaction no matter where they are in life.

Responsibilities:

- Doing Customer interaction and distributing the task among the team members.
- Done exploratory analysis and took inferences by visualization the data.
- Improvised the quality of data by removing inconsistent data, missing values and label the data using feature scaling.
- Removed insignificant variables using Dimensionality reduction technique.
- Performed Random LightFM to predict the model.
- Deployed into Ducker with the help of flask.

Wells Fargo Value prediction for each potential customer: According to Epsilon research, 80% of customers are more likely to do business with you if you provide personalized service. Banking is no exception.

The digitalization of everyday lives means that customers expect services to be delivered in a personalized and timely manner and often before they´ve even realized they need the service. Wells Fargo Group aims to go a step beyond recognizing that there is a need to provide a customer a financial service and intends to determine the amount or value of the customer's transaction. This means anticipating customer needs in a more concrete, but also simple and personal way. With so many choices for financial services, this need is greater now than ever before.

Here, Wells Fargo Group is asking us to help them identify the value of transactions for each potential customer. This is a first step that Wells Fargo needs to nail in order to personalize their services at scale.
Responsibilities:

- Doing Customer interaction and understanding the requirement.
- Done exploratory analysis and took inferences by visualization the data.
- Improvised the quality of data by removing inconsistent data, missing values and label the data using feature scaling.
- Removed insignificant variables using Dimensionality reduction technique.
- Performed Ensemble learning techniques XGBoost, LightGBM and Random Forest to predict the model.
- Deployed into Ducker with the help of flask.

Wells Fargo Customer Satisfaction: From frontline support teams to C-suites, customer satisfaction is a key measure of success. Unhappy customers don't stick around. What's more, unhappy customers rarely voice their dissatisfaction before leaving.

Wells Fargo Bank is asking to help them identify dissatisfied customers early in their relationship. Doing so would allow Wells Fargo to take proactive steps to improve a customer's happiness before it's too late.

In this project, we'll work with hundreds of anonymized features to predict if a customer is satisfied or dissatisfied with their banking experience.

Responsibilities:

- Doing Customer interaction and understanding the requirement.
- Done exploratory analysis and took inferences by visualization the data.
- Improvised the quality of data by removing inconsistent data, missing values and label the data using feature scaling.
- Removed insignificant variables using Dimensionality reduction technique.
- Performed Ensemble learning techniques XGBoost, LightGBM and Random Forest to predict the model.
- Deployed into Ducker with the help of flask.

Wells Fargo Customer Transaction Prediction: At Wells Fargo our mission is to help people and businesses prosper. We are always looking for ways to help our customers understand their financial health and identify which products and services might help them achieve their monetary goals.

Our data science team is continually challenging our machine learning algorithms to identify new ways to solve our most common challenge, binary classification problems such as: is a customer satisfied? Will a customer buy this product? Can a customer pay this loan?

In this challenge, we identifying which customers will make a specific transaction in the future, irrespective of the amount of money transacted.
Responsibilities:

- Doing Customer interaction and understanding the requirement.
- Explored the data using different visualization techniques
- Done exploratory analysis and took inferences by visualization the data.
- Improvised the quality of data by removing inconsistent data, missing values and label the data using feature scaling.
- Removed insignificant variables using Dimensionality reduction technique.
- Performed naive bayes classifier to predict the model.
- Deployed into Ducker with the help of flask.

Loan Defaulter prediction: Built a model to predict whether the banking customer will repay the loan amount or not.
Responsibilities:

- Done exploratory analysis and took inferences by visualization the data.
- Improvised the quality of data by removing inconsistent data, missing values, outliers.
- Removed insignificant variables using Dimensionality reduction technique.
- Optimized the accuracy parameters with 82% of accuracy.
- Used algorithms like KNN, NAÏVE BAYES and obtained optimal sensitivity and specificity with logistic regression.
- Deployed into Ducker with the help of flask.

Project	:	**CMSO (Commercial Managed Service Outsourcing)**
Client	:	**Shire**
Domain	:	**pharmaceutical**
Duration	:	**AUG 2014 – MAR 2015**
Technology	:	**Informatica PowerCenter 9.1, Netezza, UC4 & UNIX (putty)**
Role	:	**Software Engineer**

Shire Plc is an Irish-headquartered global specialty biopharmaceutical company. Basically we are working with products commonly known as drugs or medicines. Sales Data is logically organized information about an executed sales transaction. A sales transaction is executed when the order has been filled and the customer has been billed for the product. We are getting the sales transaction information daily, weekly, monthly, quarterly basics from different vendors and finally we are loading data into EDWARD then publishing into mart.

Responsibilities:

- Supporting onsite and offshore development team and helping them conceptualize the business processes and Preparing SQL statement for quality check.
- Created Complex Mapping and tuned them for better performance.
- Scheduling the workflow and running the workflow using third party tool UC4.

Project	:	**Single Master Data Solution-Customer Outbound Interface**
Client	:	**Maersk Line**
Domain	:	**Transportation & Logistics**
Duration	:	**September 2013– August 2014**
Technology	:	**Informatica PowerCenter 9.1, Oracle 11g & UNIX (putty)**
Role	:	**Software Developer**

This project is based on the concept of 'Single view of data'. Maersk has Very critical data related to geography, vendor, taxation, customers. For the regulatory purposes, the data need to be stored as one place rather than being scattered across their transactional systems. For every subject area, an ORS (Operational Reference Store) is defined and all these reference stores constitute an MDM Hub. All the MDM hub base objects and staging tables are loaded using Informatica MDM product.

Responsibilities:

- Supporting onsite and offshore development team and helping them conceptualize the business processes.
- Experienced with coordinating cross-functional teams, project management and presenting technical ideas to diverse groups & Proven ability to implement technology-based solutions for business problems and Created Complex Mapping and tuned them for better performance.
- Used the Informatica Designer to develop processes for data profiling, data extracting, cleansing, data scrubbing, data transforming, integrating, and loading data into the work area (staging).

Project	:	**WellPoint- Information Management system**
Client	:	**WellPoint (Anthem) Inc**
Duration	:	**MAY 2011 – AUG 2013**
Technology	:	**Informatica PowerCenter 9.1, UNIX(Reflection) & Teradata V2R6**
Role	:	**Software Developer**

WellPoint, Inc. is one of the largest health benefits companies in the United States. Through its networks nationwide, the company delivers a number of leading health benefit solutions through a broad portfolio of integrated health care plans and related services, along with a wide range of specialty products such as life and disability insurance benefits, dental, vision, behavioral health benefit services, as well as long term care insurance and flexible spending accounts.

Responsibilities:

- Worked in data Extraction, Transformation and Loading from source to target system using Tera data utilities like BTEQ, FAST Load scripts.
- Extensively used ETL to load data from Oracle and Flat files to Data Warehouse.
- Experienced with coordinating cross-functional teams, project management and presenting technical ideas to diverse groups & proven ability to implement technology-based solutions for business problems.

Declaration:

I hereby declare that the details furnished above are true to the best of my knowledge.

Date: Hrusikesh Panda

Place: