B AKHIL KUMAR

Hyderabad, IN | +918977665957 | akhilkumarb1997@gmail.com | Linkedin |Github | Portfolio

SUMMARY

Data Scientist with around 2 years of experience in using data analysis and machine learning techniques to solve business problems. Skilled in developing predictive models, conducting statistical analysis, and visualizing data using tools such as Python and R. Experienced in collaborating with cross-functional teams to identify business needs and deliver actionable insights.

EDUCATION

Mahaveer Institute of Science & Technology Hyderabad, IN
Bachelor of Technology September 2014 - October 2018
Major in Electronics and Communication Engineering

Siddhartha Junior College Hyderabad, IN
Math,Physics,Chemistry July 2012 - July 2014

EXPERIENCE

Randstad
Client: Institute for Development and Research in Banking Technology(IDRBT) Hyderabad, IN
Data Scientist & Research Associate September 2021 – Present

- Crafted **8 KNIME** workflows using **LIME**, **SHAP**, and Counterfactual methods to explain black box models for Analytical Customer Relationship Management in **Banking** and **Finance**
- Conducted research on **Variational Autoencoder** which led to development of **Chaotic Variational Autoencoder** based **One Class Classifier** for Insurance **Fraud Detection**.
- **Guided over 10 practical lab sessions** for banking professionals visiting the IDRBT for various programs.
- **Acquired, Cleaned, and Preprocessed over 15 datasets** using **machine learning** methods to gain insights.
- Collaborated with a leadership team of 4 to identify relevant questions and determine the best methods of collection, and research to increase knowledge and provide valuable contributions.
- Completed more than 20 data models and conducted analysis to produce reports that showcase the outcomes and valuable insights obtained.
- Conducted training for a group of 9 individuals from various banks on the comprehensive implementation of AI/ML models.
- Deployed 3 **End to End Machine Learning** applications using **Docker, Kubernetes and MLflow**
- Facilitated the "Advance AI/ML for Banks" program for a group of 6 individuals from diverse banking institutions.

GlobalLogic Hyderabad, IN
Associate Analyst April 2021 – September 2021

- Worked on **Google Waymo project** to tag images and videos for self-driving cars using 4 variations of cloud compute tools.
- **Presented 20+ findings and insights** to senior management to establish best practices and guide analysis into action and results.

DevOps Enabler&Co Bengaluru, IN
DevOps Engineer Intern January 2020 – June 2020

- **Designed and built 2 prototypes for CI/CD** for a website using **GCP, Git, Docker, Kubernetes, Jenkins** and documented project design for reference and future use cases.

PROJECTS

Explainable AI for Analytical Customer Relationship Management in Banking and Finance: (IDRBT)

- 8 KNIME processes were developed to learn a black box model using a generalized linear model and to use LIME/SHAP to describe the model's choices.
- Technologies/Algorithms used: **LIME,SHAP, KNIME Analytics Platform,Generalized Linear Model,RandomForest**

Chaotic Variational Auto Encoder based One Class Classifier for Insurance Fraud Detection: (IDRBT)

- Here, we employed the logistic chaotic map to generate random noise in the latent space. The effectiveness of **C-VAE** is demonstrated on the health insurance fraud and auto insurance datasets. We considered vanilla VariationalAutoEncoder (VAE) as the baseline.
- It is observed that **C-VAE outperformed VAE** in both datasets. C-VAE achieved a classification rate of 77.9% and 87.25% in health and automobile insurance datasets respectively.
- Further, the t-test conducted at 1% level of significance and 18 degrees of freedom infers that C-VAE is statistically significant than the VAE.

SKILLS

Data Science Frameworks and Libraries: Scikit-learn, NumPy, SciPy, Plotly, Pandas, Matplotlib, Tensorflow, Keras, Pytorch, H2O, MLflow, DVC, Evidently AI
Programming languages and Big Data: Python, R, Hadoop, Apache Spark, Yarn Cluster Setup
Web Development: HTML/CSS, Flask, FastAPI, Gradio
Cloud and Databases:GCP, Linux, MySQL
Machine Learning/Deep Learning Techniques: Linear Regression, Logistic Regression, Support Vector Machine, Ensemble Trees, RandomForest, A/B Testing, Data Mining, Predictive Modeling/Analytics, Statistical Analysis, Anomaly/Outlier Detection, Customer Segmentation Analysis, Artificial Neural Networks(ANN), Convolution Neural Networks(CNN), Recurrent Neural Networks(RNN), Long Short term Memory(LSTM), Variational Autoencoder(VAE), Generative Adversarial Networks(GAN)
NLP Techniques: BoW, Word2Vec, Doc2Vec, TF-IDF, BERT, NLTK, Spacy, Sentimental Analysis
CI/CD: GIT, Jenkins, Docker, Kubernetes, Github, Gitlab
Tools and IDE: KNIME Visual Analytics, Jupyter Notebook, Google Colab, Visual Studio Code

CERTIFICATIONS

- Diploma in Data Science by HSAI

ACCOMPLISHMENTS

- Got awarded 3 times as "**Alpha Player**" in **Globallogic** in the period of 6 months for being a top and accurate contributor for the project.