DILEEP SAI ELLANKI

3209 E 10TH ST APT H16 BLOOMINGTON , INDIANA, 47408

dellanki@iu.edu • (+1) 312 478 0477 • www.linkedin.com/in/dileep-sai-0464

EDUCATION

Indiana University, Bloomington, IN, United States   *August* 2022- June 2024

Master of Science in Data Science
Relevant Coursework: Data Mining, Data Visualization, Computational Linguistics & Natural Language
Processing, Artificial life in Virtual reality, Building Virtual Worlds, Statistics

SKILLS

- **Key Skills:** Data Cleaning, Data Preprocessing, Data Visualization, Predictive Analysis.
- **Technical Skills:** C Programming, Data Structures, Object-oriented Programming, Python, R-Language MATLAB, SQL, Docker, SAP BI, Tableau, Pyspark, Docker.
- **Packages:** NumPy, Pandas, Scikit-Learn, OpenCV, Seaborn, Matplotlib, NLTK.
- **Machine Learning Algorithms**: Univariate linear regression, Multivariate linear regression, Logistic Regression, K-Nearest Neighbor, Support Vector Machines, Naive-Bayes, Decision Tree, Random Forest, xgboost.
- **Deep Learning Algorithms**: Convolution Neural Networks, Generative Adversarial Networks.
- **Cloud deployment:** Google Cloud, AWS

PROFESSIONAL EXPERIENCE

Cognizant Technological Solutions, Chennai, TN, IN   July 2021 – July 2022

Data Analyst

- Learned concepts Validated and ensured the quality and reliability of over 1000 reports and datasets by analyzing data with SQL and Excel and discovering and resolving conflicts.
- Conducted exploratory data analysis (EDA) using Python and SQL, identifying trends, patterns, and anomalies in large datasets to provide insights to stakeholders.
- Developed and presented reports and visualizations to stakeholders to communicate insights and recommendations using tools such as Tableau in a clear and concise manner.

RESEARCH PUBLICATIONS

Twitter Sentiment Analysis Using Ensemble Techniques, IRJMETS,   January 2022

- This Developed a Twitter Sentiment Analysis model using ensemble techniques, including bagging and boosting algorithms, to classify tweets as positive, negative, or neutral with 85% accuracy.
- Preprocessed Twitter data using natural language processing techniques such as tokenization, stop word removal, and stemming to improve model performance.
- Presented findings to the Journal in a clear and concise manner, highlighting key insights and recommendations for future work. Stayed up-to-date on the latest developments in machine learning and natural language processing, continually improving and expanding the capabilities of the model.

PROJECTS

Speech Emotion Recognition System   October 2022 – December 2022

- Developed an Developed a Speech Emotion Recognition System using machine learning algorithms, including deep learning neural networks, to accurately classify emotions in audio samples with 75% accuracy.
- Trained machine learning models such as Random Forest, Support Vector Machines (SVM), and Convolutional Neural Networks (CNNs) to classify emotions in audio samples.
- Tested the Speech Emotion Recognition System on both male and female audio samples separately, achieving 75% accuracy on male audio and 85% accuracy on female audio. Implemented data cleaning and optimization techniques, resulting in a 30% reduction in processing time.

Airbnb Recommendation System   October 2022 – December 2022

- Developed an Airbnb Recommendation System using collaborative filtering and machine learning algorithms to show users the most suitable accommodations at a prime location within their budget, with all the facilities they need.
- Conducted data exploration, preprocessed and cleaned large datasets of Airbnb listings and user data using tools such as Python and SQL to extract and transform relevant features from the Airbnb dataset.
- Built a scalable and fault-tolerant system architecture using technologies such as Apache Spark and AWS.
- Implemented and fine-tuned collaborative filtering algorithms, such as user-based and item-based recommendations, to personalize the recommendations based on user preferences and past behaviors.
- Validated the model performance using metrics such as recall, precision, F1-score, and the area under the receiver operating characteristic (ROC) curve.