GOURI BENNI

: msoneai.com : linkedin.com/in/gouri-benni/ | : github.com/gouribenni| : gouri.benni@gmail.com | : (408)-690-4151

Summary

- Data Analyst with a focus on leveraging advanced analytics to drive operational improvements and enhance customer satisfaction. Skilled in the development and optimization of ETL workflows, adept at orchestrating sophisticated cloud-based data architectures, and proficient in distilling complex data into actionable insights.
- Solid foundation in SQL, Python, and data visualization, demonstrating a track record of refining processes, augmenting system performance, and leading cross-functional initiatives. Aligning analytical practices with strategic business goals, while committed to fostering continuous improvement in process optimization, product development, and the advancement of user experience.

Education

San Jose State University, California, USA Aug 2022 –Present

Master of Science (M.S) in Data Analytics GPA: 3.4

- Coursework: Data Visualization, Data Mining and Analytics, Databases, Stats, Machine Learning, Deep Learning, Data Analytics, Management.

Gogte Institute of Technology, Karnataka, India Aug 2016 – Sep 2020

Bachelor of Engineering (B.E) in Computer Science Engineering GPA: 3.5

- Coursework: Python, Databases Data Science, Data Structures and Algorithms, Software Engineering, Cloud Computing, Artificial Intelligence.

Work Experience

Norfolk Southern Corp May 2023 – Aug 2023

Data Analyst Georgia, USA

- Conducted advanced data analytics to analyze train scheduling and logistics, using SQL for time series analysis; identified patterns leading to a 25% reduction in train delays and enhanced customer satisfaction.
- Developed and optimized ETL processes using Python with AWS Glue, improving large-scale data handling efficiency; managed data within AWS RedShift for better operational decision-making.
- Created Tableau dashboards for train delay monitoring, focusing on departure, arrival times and delay causes, leading to improved operational efficiency and reliability.
- Facilitated data-driven operational improvements by engaging in strategic communication and collaboration with operational teams to mitigate root causes of train delays and meticulously monitoring and tracking key performance indicators related to train schedules.

ADP (Automatic Data Processing) Apr 2021 – Jun 2022

Data Analyst Hyderabad, India

- Participated in strategic analytics and data engineering projects that amplified payroll processing efficiency by 20% through targeted statistical analysis and A/B testing, while simultaneously enhancing SQL query performance by 40% through the optimization of over 30 ETL scripts and reducing payroll discrepancies with a co-developed regression model.
- Engineered complex data visualization charts like t-sne and Sankey using plotly library in Python for real-time payroll metrics monitoring, instrumental in cutting processing cycle times by 15% and elevating client satisfaction levels.
- Orchestrated a project to refine payroll workflows using JIRA and Gantt charts, which delivered a 30% boost in issue resolution and a 25% rise in processing productivity.
- Implemented data governance and enforced quality control protocols, securing payroll data accuracy and compliance with financial regulations.
- Synthesized business intelligence with payroll strategies to meet client requirements, led cross-functional teams for systemic improvements, and integrated Agile methodologies to advance payroll data pipeline development and testing, fostering an environment of continuous innovation and adaptability.

Projects

Job Transition Pathway:

- Spearheaded a data science project to create an AI-powered job transition pathway using generative LSTM models, synthesizing data from multiple sources including Indeed for jobs, Udemy for courses, and EMSI for skills.
- Conducted thorough data cleaning and processing, employing pre-built NLP tools like SpaCy for data normalization, and developed a fuzzy matching algorithm to associate job progression stages with corresponding skill sets.
- Leveraged BERT embeddings within a Seq2Seq LSTM framework to capture contextual nuances in job descriptions, ensuring the model's deep linguistic comprehension for accurate job and skill predictions.
- Designed and evaluated an Seq2seq model with encoder adn decoder that predicts individualized career progression paths and provides tailored course recommendations, utilizing exploratory data analysis (EDA) for synthetic data generation to enhance model training.
- Validated the model using a separate dataset to ensure generalizability, achieving a BLEU score of 51.43% for job predictions and 68.84% for skill predictions, reflecting the model's adept training assimilation and its proficiency in predicting career progression pathways.

Ranked Stack Overflow: Mathematics and Statistical Analysis

- Conducted comprehensive NLP analysis on over 370,000 Stack Overflow threads, developing an advanced ranking model. This model significantly boosted answer retrieval accuracy by 64%, showcasing expertise in applying NLP for real-world data challenges.
- Implemented machine learning classifiers like Random Forest and SVM for content tagging, achieving 96% precision. Enhanced the relevance of solutions through the deployment of 'all-MiniLM-L6-v2' and 'all-mpnet-base-v2' transformer models, reflecting a deep understanding of AI technologies in semantic analysis.
- Pioneered a personalized ranking algorithm, improving NDCG scores by 21%. This innovation, combined with a 30% reduction in query resolution time, significantly enhanced user experience on educational forums, demonstrating a balanced focus on technical and user-centric design.

Skills

- **Programming Languages:** Python (Pandas, NumPy, Matplotlib, Seaborn, SciPy), R, C/C++, HTML, CSS
- **Databases:** MySQL, NoSQL, ETL, MS SQL Server, Nested Queries
- **Tools:** Advanced Excel, Tableau, Jupyter Notebook, GIT, Airflow, AWS, Google Data Studio, PowerBI
- **Machine Learning:** Clustering, Business Prediction, Timeseries Forecasting, Classification and Regression.
- **Project Management tools:** JIRA, Agile (Scrum) methodologies, Gantt chart.