Yashkumar Chandwani Data Scientist

✉ chandwaniyashkumar7@gmail.com 📞 (817) 677-0173 📍 Arlington, TX in LinkedIn ○ GitHub

Profile

Motivated and Analytical Data Scientist with 2+ years of experience with a strong foundation in Statistics, Machine Learning, and Programming. Equipped with hands-on experience in data analysis, visualization, and model development with Python and SQL. Committed to applying data-driven insights to solve complex problems and make informed business decisions.

Skills

Languages and IDE: Python, SQL, HTML, CSS, C, VBA, Google Colab, Spyder, Pycharm, Jupyter Lab, Visual Studio, R

Cloud & Big Data: Microsoft Azure (Databricks, ADF, Data Lakes, Delta Lakes, Azure Blob Storage, Azure Synapse), AWS (EC2, S3, Redshift, Glue, EMR, Lambda, Dynamo DB, Snowflake), Hadoop, HDFS, Apache Spark, Apache Airflow, PySpark

Frameworks & Tools: Tableau, Power BI, Django, Flask, GenAI (Vector DB, Llama Index, Lang Chain, Vertex AI), Numpy, Pandas, Matplotlib, ArcGIS, Alteryx, Docker, Kubernetes, Informatica, IICS, Microsoft Excel, PowerPoint, SAP

Machine Learning and Deep Learning: Random Forest, Decision Tree, Scikit Learn, XGBoost, CNN, RNN, LSTM, Spacy, LLM, Tensorflow, Keras, NLTK, Scikit Learn, Time Series, Statistics (A/B Testing, Probability), Statistical Analysis

Professional Experience

Data Scientist, *Capgemini*   01/2024 – Present │ Remote, USA
- Spearheaded the acquisition, preprocessing, and analysis of 50+ datasets (10 TB) using Python and SQL, producing 100+ charts and graphs (Matplotlib, Seaborn, Power BI) to communicate insights.
- Collaborated with senior data scientists to develop predictive models leveraging regression, decision trees, and random forests, resulting in a 20% improvement in accuracy through feature engineering.
- Orchestrated model evaluation using cross-validation techniques, achieving average metrics of 90% accuracy, 85% precision, 88% recall, and 87% F1 score.
- Integrated deep learning techniques such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) into predictive modeling processes, resulting in a further 10% improvement in accuracy over traditional methods.
- Created ML pipelines using LLMs for auto data preprocessing, model training, prediction tasks, sentiment analysis, and text generation, enhancing NLP capabilities, enriching insights from unstructured data, and improving 30% of operations.

Data Scientist, *Maxgen Technology*   01/2021 – 12/2022 │ Remote, India
- Orchestrated the development and deployment of a high-impact classification model using Python and relevant libraries, resulting in a 15% revenue surge and enhancing customer engagement by 10%.
- Spearheaded predictive model development with Python, reducing customer churn rate by 15% by using sophisticated analyses of demographics, transaction history, and customer behavior through libraries like Pandas and Scikit-learn.
- Directed end-to-end data processes encompassing extraction and preprocessing using Python, showcasing proficiency in statistical analysis to unearth actionable insights.
- Applied advanced techniques including NLP and sentiment analysis with libraries such as NLTK and SpaCy to derive meaningful insights from customer feedback, facilitating clustering analysis for trend identification and strategy.
- Utilized Tableau for data visualization, creating interactive dashboards that provided clear and actionable insights.
- Leveraged pivot tables, array functions, and Python automation scripts to revolutionize data analysis, automate processing tasks, and deliver a 30% reduction in analysis time and a 25% boost in overall efficiency.

Education

Master of Science, *University of Texas*   01/2023 – 12/2024 │ Arlington, USA
Computer Science

Bachelor of Science, *Gujarat Technological University*   06/2018 – 06/2022 │ Ahmedabad, India
 Information Technology

Projects

Chatbot, *Python, NLP, TensorFlow, Chatbot Frameworks and RESTful APIs*   04/2023
- Collaborated on the development of an intelligent chatbot designed to provide personalized recommendations and answer user queries in real time.
- The chatbot effectively enhanced user engagement and streamlined customer interactions.

AI-Math-Notes, *Python, OpenAI, Tkinter, and Canvas*   03/2023
- Developed AI-Math-Notes, an educational tool that integrates OpenAI's model to generate explanations and solutions for mathematical problems, enhancing learning with dynamic content rendering and user interaction on a Canvas interface.
- Designed and implemented a user-friendly interface with Tkinter, allowing seamless navigation and intuitive interaction for students and educators alike, promoting accessibility and usability in STEM education environments.

Face Recognition Attendance System,   02/2023
Python, TensorFlow, OpenCV, Dlib, and Facial Recognition Algorithms
- Developed a system that utilizes deep learning techniques to automate attendance tracking.
- Contributed to the development of a user-friendly interface for instructors to manage attendance records efficiently