

Tharun Kumar Bottlapally Data Scientist

Tharunreddy0104@gmail.com | +1 (980) 382-7693 | Charlotte, NC | LinkedIn

Summary

Accomplished Data Scientist with 4 years of experience excelling in the utilization of data to inform strategic business decisions. Proficient in employing advanced statistical analysis and cutting-edge machine learning techniques to derive actionable insights from intricate datasets. Adept at designing and implementing predictive models, conducting comprehensive data exploration, and creating scalable data pipelines.

Technical Skills

- **Languages:** Python, R Studio, SQL, Java, PHP, C, C++, JavaScript, HTML, CSS.
- **Data Science:** Machine Learning Algorithms, Deep Learning, NLP, Regression, Classification, Clustering, Recommendation, Neural Networks, TensorFlow, BERT, ARIMA, TensorFlow, Keras, Sci-kit Learn, Topic modeling, FRCNN, YOLO, ResNet50, LSTM, K-mean Cluster, GAN, CNN, RNN, FFNN, NTLK.
- **Database:** MySQL, MongoDB, PostgreSQL, NoSQL, Amazon Redshift, MS SQL Server, Snowflake.
- **Data Processing & Streaming:** NumPy, Pandas, Apache Spark/PySpark, Hadoop, Apache Hive, Apache Kafka, Matplotlib, Seaborn, Altair, Plotly, Big Data, Big Query, Pycharm, Airflow, PyTorch, XGBoost, SciPy, Spacy.
- **Others & Tools:** Power BI, Tableau, Git, GitHub, Bitbucket, AWS, Microsoft Azure, Django, Streamlit, Flask, Statistics.

Professional Experience

Data Scientist, *BatteryXchange*

07/2024 – Present | Charlotte, USA

- Developed and implemented a comprehensive IT Disaster Recovery Plan, reducing potential downtime by 50%, ensuring 99.9% data integrity, and enhancing business continuity through detailed risk management, backup strategies, emergency procedures, and vendor security management.
- Designed and deployed dynamic data pipelines using Apache Airflow, integrating data from cloud MySQL, Postgres, and MongoDB databases, reducing reporting time by 40%, enhancing data accuracy by 30%, and improving data accessibility for over 15 clients, leading to more efficient decision-making processes.
- Developed and deployed machine learning models using scikit-learn and TensorFlow for predictive analytics, resulting in a 25% increase in sales conversion rates and a 15% reduction in customer churn.
- Collaborated with cross-functional teams to define and implement data governance policies, ensuring data quality, security, and compliance with industry regulations such as GDPR and HIPAA.

Data Scientist, *Wells Fargo*

01/2024 – 05/2024 | Remote, USA

- Collected diverse financial data sources, including stock market data, economic indicators, and news sentiment, leveraging AWS data storage solutions such as Amazon S3 to store over 10TB of data for efficient retrieval and analysis.
- Preprocessed data using Python libraries like Pandas and NumPy to handle missing values and outliers, ensuring data quality and creating a structured dataset of 1 million+ records for training and testing machine learning models.
- Utilized machine learning algorithms including Random Forest, Gradient Boosting, and Support Vector Machines with AWS SageMaker to predict investment outcomes, achieving a 92% accuracy rate.
- Implemented deep learning models with TensorFlow/Keras for complex predictions, experimenting with NLP and LSTM networks for time-series analysis, improving accuracy by 15%.
- Developed a user-friendly web interface or dashboard for analysts to input parameters and view investment predictions, enhancing decision-making processes and reducing analysis time by 40%.
- Deployed scalable and reliable investment prediction models using AWS Lambda, Fargate, and ECS, with CloudFormation for infrastructure management and CloudWatch for monitoring and logging.

Data Scientist, *ADP Pvt. Ltd*

10/2021 – 07/2022 | Hyderabad, India

- Conducted extensive statistical analysis and time series forecasting on sales and inventory data, resulting in a 20% improvement in predictive model accuracy, which facilitated more informed decision-making regarding stock levels and promotions.
- Developed and implemented advanced anomaly detection algorithms to identify irregularities in sales patterns, reducing false positives by 25%. This improvement significantly enhanced system stability and ensured data integrity.
- Architected and deployed scalable machine learning models in Python using Scikit-learn, Pandas, and NumPy, resulting in a 30% reduction in model training and inference times for faster insights and actions.

- Utilized advanced NLP techniques with PyTorch, TensorFlow, and Hugging Face Transformers to enhance text classification accuracy by 15%, facilitating real-time sentiment analysis that guided marketing strategies.
- Optimized data processing pipelines in Jupyter Notebook and Databricks using Pandas and PyArrow, achieving a 35% reduction in ETL processing times and improving overall data handling efficiency.
- Developed innovative prompt engineering strategies using LangChain, boosting AI-driven customer interactions by 20% and enhancing response accuracy in chatbot applications on the client's website.
- Successfully deployed machine learning models on AWS cloud services, including SageMaker, Lambda, Glue, and EC2, resulting in a 40% increase in operational efficiency and a 50% reduction in deployment time across multiple environments, ensuring seamless integration into the client's infrastructure.

Data Scientist, Cognizant Technology

12/2019 – 08/2021 | Hyderabad, India

- Spearheaded the application of advanced data analysis techniques, utilizing regular expressions (regex), spaCy, and Yake, to extract critical pricing data and pinpoint risk-related keywords.
- Realized a notable 40% increase in profit margins through the application of data-driven insights, guiding strategic trading decisions.
- Employed Azure Sentiment Analysis and Opinion Mining to evaluate trader sentiment and analyze the impact of external factors on trading. Resulted in an impressive 15% enhancement in decision accuracy.
- Collaborated closely with a dynamic team to conceive and implement robust risk management strategies, leading to a commendable 25% reduction in overall portfolio risk.
- Meticulously managed and maintained a SQL database for comprehensive tracking of trade signals, executed trades, and profitability. This initiative resulted in a significant 30% reduction in data retrieval times.
- Successfully deployed the project on AWS EC2, enhancing scalability and streamlining data management. Utilized AWS RDS for a high-performance relational database, achieving a remarkable 40% improvement in data access speed.

Education

Master of Science, The University of North Carolina at Charlotte

08/2022 - 05/2024 | Charlotte, USA

Data Science and Business Analytics

Bachelor of Technology, JNTU

08/2016 - 09/2020 | Hyderabad, India

Computer Science and Engineering

Projects

JARVIS: The Data Mining Robot

- Built a virtual robot, JARVIS, to scrape Corporate Action data from 2 websites. Coordinated users with stock prices based on Corporate Action data.

Brazilian E-Commerce Data Analytics and Reporting System

- Led the successful development of a MySQL-powered data analytics platform, expertly handling a vast dataset comprising more than 10,000 data points. Utilized advanced SQL queries and Key Performance Indicators to extract valuable insights, enabling data-driven decision-making and facilitating business growth.

Kaggle Machine Learning and Data Science Survey Analysis

- Led the development and execution of a comprehensive data science survey analysis, uncovering impactful insights on demographics, education, and industry trends from 2018 to 2021; informed strategic initiatives that drove revenue growth, optimized marketing campaigns, and improved customer satisfaction.

Supermarket Billing System

- Engineered a robust software solution utilizing JAVA and SQL Server, revolutionizing the billing process in supermarkets; achieved significant time savings by generating bills quickly, while seamlessly storing details in a centralized database.

Accomplishments

- Published research paper on Job Shifting Prediction and Analysis Using Machine Learning ([Link](#))
- Smart India Hackathon March 2019 Grand Finalist
- Participated in the 2023 Carolina Hurricanes Business Analytics Challenge.
- Truist 2023 Data Modelling Competition Grand Finalist.