NIKHIL R PALLEPATI

(571) 352 – 0572 | nikhilreddyp29@gmail.com | LinkedIn | GitHub | Kaggle

SUMMARY: With over 4 years of experience in Data Science, have honed my expertise in Machine Learning, Natural Language Processing (NLP), and Knowledge Graph design, aligning well with the Applied Scientist role. My core strength lies in translating complex data into actionable insights, developing algorithms, and optimizing models to address real-world challenges.

EDUCATION

CARNEGIE MELLON UNIVERSITY, Pittsburgh, PA August 2023

Master of Information Systems Management, Business Intelligence and Data Analytics

***Relevant Coursework*:** Machine Learning, Intermediate Statistics, Unstructured Data Analytics, Econometrics, Distributed Systems, Database Management, Time Series Forecasting, Managing Analytics Projects, Advanced Business Analytics, Agile Methodologies.

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY, Hyderabad, India April 2018

Bachelor of Technology, Electronics and Communication Engineering

WORK EXPERIENCE

cPacket Networks (Capstone Project) Pittsburgh, PA

Data Scientist *May 2023 – August 2023*

- Spearheaded design and deployment of **Knowledge Graphs (KG)** using community detection algorithms, resulting in enhanced anomaly classification and strategic pathway for future graph-based ML explorations.
- Optimized the extraction, transformation, and load **(ETL)** process from cStor to KG and facilitated extraction of complex network insights. This resulted in a **30-hour** network downtime reduction, reflecting tangible increase in operational efficiency.

Tata Consultancy Services – Digital Hyderabad, India

Systems Engineer (Machine Learning Developer), Client: Reserve Bank of India. *March 2020 – July 2022*

- Analyzed major Indian Realtors Data, derived the House Pricing Index, and predicted trends for states and major cities of India.
- Built Natural Language Processing (NLP) model for automated product categorization, resulting in **80%** improvement in efficiency in calculating Consumer Price Index over time.
- Developed and deployed a Sentiment Analysis model, utilizing dataset of **200,000** news articles to extract milestones discussed and overall polarity in news articles related to Reserve Bank mentions.

Assistant Systems Engineer (Analyst), Client: U.S. Insurance Major *February 2019 – March 2020*

- Designed and implemented NLP model that automated ticket prioritization of the logged incidents, boosting productivity by **65%.** Further, implemented **A/B Testing** to evaluate effectiveness of this system, which led to subsequent **10%** increase in efficiency.
- Built and deployed predictive models for marketing optimization and customer engagement, used customer survey text analytics and Exploratory Data Analysis **(EDA).** Effectively communicate insights to stakeholders, thus enhanced decision-making process.

Tata Projects Limited Raipur, India

Graduate Intern *September 2018 – January 2019*

- Part of Social Media Analytics team for Govt. of Chhattisgarh, designed data models for analytical reporting. Created interactive Power BI dashboards and ad-hoc reports to support decision-making for relevant stakeholders.

ACADEMIC PROJECTS

Data-Driven Peer-to-Peer Lending Investment Strategy Analysis *March 2023 – April 2023*

- Developed Statistical predictive models using various ML techniques on **Lending Club** dataset and determined support vector machines as the best, achieved an accuracy of **82%,** to estimate loan default probabilities and potential returns to investors.

Overcoming Biases in Toxicity Models for Inclusive Conversation *March 2023 – May 2023*

- Utilized **BERT** techniques for AI-based toxic comment detection system, achieved **80%** reduction in harmful content, minimized biases by **60%.** Implemented an AI solution, enhanced student inclusivity, well-being, and reduced moderation efforts by **70%.**

Question Answering (QA) System *February 2023* – March 2023

- Implemented data preprocessing steps and enhanced QA accuracy to **75%** through **Jaccard overlap**-based answer identification. Improved model accuracy to **77%** by fine-tuning**,** and deploying **BERT** model, with custom tokenizer, optimized gradient descent.

Movie Recommendation System *January 2023 – February 2023*

- Conducted evaluation of user behavior patterns to generate personalized recommendations by Collaborative Filtering approach and applied user-user similarity, item-item similarity, and matrix factorization algorithms.

SKILLS

- **Programming Languages** : Python, SQL, Java, and R.
- **Relevant Libraries** : NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, SpaCy, NLTK, PyTorch, TensorFlow, Huggingface.
- **Databases and Other Tools:** MySQL, MongoDB, Neo4j, Informatica, Tableau, Streamlit, Airflow, Git, Jenkins, Databricks.
- **Big Data and Cloud** : Spark (PySpark, Spark Java), Hive, Pig, Kafka, Docker, Kubernetes, AWS, Azure, GCP.