

## CHANDU PRASAD REDDY POTUKANUMA

Chicago, IL | Mobile: +1 773-879-6504 | Email: [cpotukanuma@gmail.com](mailto:cpotukanuma@gmail.com) | LinkedIn: [Chandu Reddy](#)

### SUMMARY

Highly skilled Data Engineer with over 2 years of professional experience working with cutting-edge technologies and platforms. Adept at utilizing Google Cloud Platform (GCP) tools, PySpark, and SQL to create scalable and efficient data solutions. Experienced in designing and developing ETL pipelines, optimizing workflows, and collaborating in Agile/Scrum environments. Knowledgeable in machine learning, certified in Machine Learning Foundations, and proficient in various operating systems and databases.

### TECHNICAL SKILLS

Programming Language:	Python, SQL, Scala
Big Data Tools:	Apache Spark, PySpark, Hadoop
Cloud Technologies:	Google Cloud Platform(GCP), Amazon Web Services(AWS)
Databases:	MySQL, SQL Server, Oracle, MongoDB
Version Control & CI/CD:	Git, GitHub, Jira, Jenkins
ETL Tools	Apache Airflow, Informatica
Machine Learning Tools:	Scikit-Learn, TensorFlow
Operating Systems:	Mac OS, Windows, Linux/Unix
Visualization Tools:	Tableau, Power BI

### PROFESSIONAL EXPERIENCE

#### WALMART, India

July 2021 – December 2022

##### Big Data Developer

- Leveraged GCP tools such as Big Query, Cloud Storage, and Dataflow for managing and analyzing large datasets.
- Converted PySpark scripts to SQL and developed views in GCP to improve data accessibility and performance.
- Designed and developed Airflow DAGs for orchestrating complex data workflows and ensuring data consistency.
- Built scalable ETL pipelines to ingest, transform, and load data from Archer into GCP.
- Created comprehensive documentation and one-pagers using Confluence for team-wide knowledge sharing.
- Implemented data migration strategies for moving legacy data to modern cloud-based solutions.
- Collaborated with cross-functional teams to ensure seamless integration of machine learning models into production pipelines.

### EDUCATION AND CERTIFICATIONS

#### Chicago State University – Illinois, USA

Master of Science in Computer Science

#### JNTU Anantapur -Kakinada, India

Bachelor of Technology: Electronics and Communication Engineering (ECE)

**Machine Learning Foundations Certification** – AWS Academy

**Data Engineer Certification** – Coursera

### PROJECTS

#### Fake Tweet Detection Model

- Built and deployed a machine learning model using CNN to classify bot-generated tweets in Django.
- Conducted extensive preprocessing of text data, including tokenization, stemming, and removal of stop words, to improve model input quality.
- Implemented a real-time API using Django REST Framework for on-the-fly predictions.
- Designed an intuitive front-end interface in Django for end-users to upload tweets and view predictions.
- Conducted hyperparameter tuning to enhance model accuracy, achieving a classification accuracy of 92% on test data.

## Wine Quality Analysis

- Conducted data analysis and visualization on the UCI Wine Quality dataset using Pandas and Matplotlib.
- Developed a machine learning model to predict wine quality using Random Forest and Gradient Boosting algorithms.
- Performed extensive feature engineering, including correlation analysis and scaling of numeric features.
- Created comprehensive visualizations to explore the relationships between wine attributes and quality ratings.
- Achieved an accuracy of 85% on test data and deployed the model using Flask for end-user interaction.

## KEY STRENGTHS

---

- Proficient in Python and SQL for data manipulation and analysis.
- Skilled in cloud technologies and data pipeline automation.
- Strong communicator with a proven ability to document and share knowledge effectively.
- Agile and adaptable, with hands-on experience in fast-paced, collaborative environment.