Venkata Chirravuri

4259 Cotswolds Hill Ln, Fairfax, VA | +1 (571)-519-8090 | venkatachirravuri95@gmail.com|Portfolio| GitHub| LinkedIn

EDUCATION

George Mason University, Fairfax, VA **August 2021 - May 2023**
Masters of Science in Computer Science : *3.4/4.0*

Mahindra University, Hyderabad, India **August 2014 - June 2018**
Bachelor of Technology in Computer Science : *7.29/10.0*

SKILLS

Programming Languages:	Python, JavaScript, SQL, HTML
Databases:	MySQL, PostgreSQL, MongoDB, Snowflake, Cassandra
Big data Technologies:	Hadoop, HDFS-MapReduce, Hive, Spark, PySpark, Spark Streaming
Frameworks:	CSS, Apache Kafka, Airflow, Django, Flask, Fast API, Rest API AngularJS
Cloud Technologies:	AWS (IAM, EC2, EMR, S3, RDS, RedShift, SNS, SQS, CloudWatch, Glue, Kinesis, Athena CloudFormation), GCP Big Query
Visualization & CI/CD:	Power BI, Matplotlib, Tableau, Docker, Git, Kubernetes
Other :	Agile Methodology, Scrum, Kanban, Cross-functional Collaboration, Sprint Planning

PROFESSIONAL EXPERIENCE

Data Engineer at Tech Mahindra, Bengaluru, India **August 2019 - June 2021**

- **United Data Migration Framework (UDMF)**:
  - Developed efficient data migration libraries leveraging Spark, resulting in a successful implementation of UDMF 2.0 for Verizon; migrated high volume data, enhancing data processing speed by 40% and reducing migration time by 50%.
  - Optimized native reusable libraries functionalities by implementing Python, Scala, and SQL techniques, resulting in a 40% reduction in processing time.
  - The framework has improved overall system performance and has demonstrated a 50% reduction in overhead costs.
- **Data Migration: Teradata - Google Big Query**:
  - Brainstormed and proposed a proof-of-concept for Tech Mahindra and other business partners, facilitating seamless data migration from a database management system to a cloud-based data warehouse.
  - Created Python scripts to automate hassle-free data migration from Teradata to Google BigQuery, resulting in a 50% reduction in data transfer time and enabling faster data analysis and reporting.
  - Planned relevant metrics, objectives and constraints for data pre-processing techniques and data pipelines.

Jr. Data Scientist at Tech Mahindra, Bengaluru, India **July 2018 - August 2019**

- **Artificial Intelligent Legal Assistant (AILA)**:
  - Collaborated in a team to develop AILA, an application to automate legal contracts management using NLP techniques.
  - Preprocessed and analyzed 7000 contracts to train and test bi-directional LSTM text classifiers.
  - Led a team, including Legal experts, analyzed strategies for rule-based NER approach to extract provisions, metadata, and risks with 94% accuracy.
  - The platform reduces turnaround time from 10 days to 7 hours and increases accuracy by minimizing manual errors.

ACADEMIC PROJECTS

Collaborative Immersive Analytics **Unity, Immersive Analytics Toolkit (IATK), Oculus Integration, C#, VRTK Tilia**

- Conducted research and redefined a collaborative 3D Virtual Reality environment for data analysis using IATK. Enabled multi-user interaction with Unity, Unity PUN, and VRTK Tilia packages, deploying the application on Oculus 2 devices.
- Coordinated the implementation of an open-source infrastructure for Multi-user XR research, Collaborative Mobile Immersive Computing, resulting in a 40% increase in research productivity and collaboration efficiency.

Car Price Prediction **PySpark, MLlib, Handy Spark**

- Took lead in designing and developing a precise approach for estimating used car prices, leveraging data from Craigslist.
- Employed PySpark & MLlib for efficient data processing, cutting processing time by 60% and boosting predictive accuracy by 45%. Created Power BI dashboards for comprehensive analysis of multiple machine learning models' results.

Student Survey Portal **HTML/CSS, JavaScript, AWS (S3, EC2), Kubernetes, Docker, Jenkins**

- Gained expertise in full-stack development, leveraging AWS services (S3, EC2, IAM) for application creation and hosting.
- Employed Docker and Kubernetes for reliable and scalable application containerization and deployment, in Rancher-managed K8 clusters. Enhanced product updates speed and reliability by setting up CI/CD pipelines with Jenkins.