JYOTHIR VENKATA PAVAN VARMA VATSAVAYI

Linkedin: https://www.linkedin.com/in/jvatsava/

Email: jvatsava@asu.edu
Mobile: +1 602-768-7446

EDUCATION

- **Masters in Computer Science** — Aug 2022 - Present
  Arizona State University — *Tempe, Arizona*
  GPA - 3.89/4.00

- **B.Tech - Computer Science and Engineering** — Aug 2018 - Sep 2022
  Anna University — *Chennai, India*
  GPA - 8.01/10.00

TECHNICAL SKILLS

Cloud Services: Azure Databricks, Azure Data Factory, Azure Synapse Analytics, Synapse Pipelines.
Frameworks and Tools: PySpark, Power BI, GitHub, Selenium, HTML, CSS, Exel.
Languages and Databases: Python, Java, SQL, Javascript, MongoDB, Cosmos DB.

PROFESSIONAL EXPERIENCE

- **Arizona State University**
  Graduate Service Assistant — CSE579 Knowledge Representation — *Jan 2023 - Present*
  - Evaluated and provided comprehensive feedback on 200 student assignments, projects, and exams focusing on Clingo and Protege, ensuring adherence to course objectives and facilitating students' understanding and proficiency in the subject matter.

- **Arizona State University**
  Teaching Assistant — CSE110 Principles of Programming in JAVA — *Aug 2022 - Dec 2022*
  - Lecturing Principles of Programming to 155 students in their Lab Sessions.
  - Handling 4 Study Hall Sessions to clear the students doubts.

- **U.N.I.Q Technologies (Intern)**
  Cloud — *Jul 2021 - Aug 2021*
  - Assisted in configuring and managing cloud environments, leveraging platforms like Azure or AWS for optimal data processing and storage. Focused on enhancing cloud resource utilization, ensuring data security, and maintaining system scalability.

RELEVANT PROJECTS

- **Azure Databricks Data Engineering Project (Formula 1 Datasets)**:
  1. Engineered a Formula 1 data analytics project utilizing Azure Databricks and Spark Core, analyzing historical data from 1950 to 2017. Successfully set up and managed a dedicated cluster in Azure Databricks, ensuring optimal data processing.
  2. Created a comprehensive and efficient data solution integrating Azure Databricks, Azure Data Lake Gen2, Data Factory, and Power BI. This system was tailored for seamless data processing, effective visualization, and secure handling of sensitive data including access tokens and secrets.
  3. Organized and processed various data formats including CSV and JSON, implementing an efficient storage system categorized into bronze, silver and gold layers. Optimized data storage and querying by converting data into Parquet and Delta formats.
  4. Implemented Delta Lake's time-travel feature for precise data version control and integrity, and autonomously developed automated data pipelines within a Lakehouse architecture, significantly enhancing data processing efficiency and reducing manual workload.

- **Azure Data Factory Project (COVID-19 Trend Analysis)** :
  1. Developed a COVID-19 trend analysis solution using Azure Data Factory, orchestrating data integration from multiple sources including HTTP clients, Azure Blob Storage, and Azure Data Lake Gen2, ensuring a seamless data management process.
  2. Managed secure data storage utilizing Blob Storage, and SQL Database, coupled with complex data transformations within Azure Data Factory through source filtering, aggregation, and joining processes.
  3. Designed and executed dynamic data pipeline workflows in Azure Data Factory autonomously, leveraging control flow activities and diverse triggers like Event, Schedule, and Tumbling Window to maintain consistent and effective data flow.
  4. Integrated Azure HDInsight and Azure Databricks for advanced data processing and analysis, creating in-depth Power BI reports for visualizing and interpreting COVID-19 trends and analytics, thereby driving insightful predictions and analyses.

- **Azure Synapse Analytics Project (NYC Taxi Data)** :
  1. Built a project utilizing Azure Synapse Analytics to process and analyze NYC Taxi Trips data, integrating components like Serverless SQL Pool, Spark Pool, and Power BI for comprehensive data analysis and visualization.
  2. Designed and implemented scalable data processing workflows using Azure Synapse Analytics Architecture, Synapse Pipelines, and Azure Data Lake Storage Gen2, ensuring efficient management and analysis of large datasets.
  3. Architected real-time analytics solutions leveraging Synapse Link for Cosmos DB and HTAP capabilities, and created dynamic, insightful reports and dashboards in Power BI.
  4. Successfully navigated challenges in data scalability and integration of various Azure services, enhancing skills in cloud-based data warehousing and real-time analytics, applicable to roles requiring expertise in data management and business intelligence.