Spoorthi Veeresh Kumar

608-440-4929 |spoorthiv20@gmail.com |linkedin.com/in/spoorthiv21 |github.com/spoorthi221 |Madison,WI(Open to Relocate)

Education

University of Wisconsin Madison   **August 2023 – May 2025**
Master of Science in Information (Data Track)STEM   *United States*

Visvesvaraya Technological University   **August 2019 – May 2023**
Bachelor of Science in Computer Science and Engineering   *Karnataka,India*

Relevant Coursework: Data Structures, Artificial intelligence/Machine learning, Relational Database Management Systems, Big Data Analytics.

Technical Skills

Programming Language: Python, SQL, R
Data Visualization Tools: Tableau, Excel, Matplotlib, Seaborn
Frameworks and Libraries: Numpy, pandas,TensorFlow, PyTorch, GitHub
Big Data Cloud: Apache Spark, Hadoop, AWS (S3, Redshift, Glue, Lambda, EMR)

Experience

Interviewer.Ai   **Jan 2024 – Present**
Data Scientist   *Remote*

- Designed and deployed machine learning models achieving 81% accuracy in image classification using ResNet-50 on a unstructured dataset of 50,000 images, improving previous model performance by 13%.
- Conducted DPIA fairness assessments across protected attributes (race, gender, age), achieving above 0.85 demographic parity and ensuring GDPR compliance.
- Delivered end-to-end reliable and scalable AI features in production.

UW - International Academic Program   **July 2024 – Present**
IAP Data Analyst Support   *Madison, WI*

- Validated and analyzed large datasets, improving data-driven decision-making and optimizing participation metrics by 12%.
- Developed Tableau dashboards with interactive filters to analyze student mobility trends and program effectiveness.

University of Wisconsin Madison   **Nov 2023 – May 2024**
Research Assistant   *Remote*

- Developed an evaluation framework for assessing ChatGPT's LLM clustering capabilities in educational contexts, achieving a 9% improvement in label specificity across 500+ document clusters
- Created interactive visualizations of research findings that were incorporated into departmental presentations and grant applications

Addicor Technologies   **Jan 2023 – June 2023**
Data Analyst Intern   *Remote*

- Optimized SQL queries and design to enhance performance, improving financial reporting efficiency by 19% and serving 7+ stakeholder departments.
- Designed and maintain models to ensure relationships between datasets.

ISans Technologies   **Sep 2022 - Feb 2023**
AWS Cloud Intern   *Bangalore*

- Designed and implemented a CI/CD pipeline using Jenkins, reducing deployment time by 10%, with a focus on cloud environments.

Projects

Scalable Supply Chain Data Pipeline using Apache Spark & AWS | *Apache Spark, AWS, PostgreSQL, Airflow, Elasticsearch*

- Ingested structured unstructured data from IoT sensors, warehouse logs, and transportation tracking APIs into AWS S3.
- Integrated Elasticsearch for complex search queries on shipment status, manufacturing logs, and failure reports.
- Designed an ETL pipeline using PySpark to clean, transform, and aggregate data, reducing processing time by 40%.

Pursuit Data Integration Pipeline | *ApacheSpark, PostgreSQL, ElasticSearch*

- Built scalable data pipelines using Apache Spark PostgreSQL to support complex entity mapping and search queries.
- Designed an ETL pipeline using PySpark to clean, transform, and aggregate data, reducing processing time by 40%.
- Proposed read/write separation strategies, table partitioning, and caching optimizations for scaling.