K. HARITHA

+91 9496823295 ◇ Ernakulam, Kerala, India
haritha.kaladharan68@gmail.com ◇ https://www.linkedin.com/in/harithakaladharan/

Enthusiastic researcher with over 6 years of experience in Artificial Intelligence and Data Science. Specializing in big data analytics, machine learning, data mining, predictive analytics, and cognitive computing. Also, equipped with 10 months of industrial experience in advanced AI technologies such as Generative AI models (e.g., Large Language Models - LLMs, Retrieval-Augmented Generation - RAG) and expertise in Big Data Technologies like Apache Spark, Hive, Hadoop, as well as web development tools like Django. Demonstrated success in developing and implementing ML/DL models to address complex challenges and drive innovation.

EDUCATION

Doctor of Philosophy (ABD), Cochin University of Science and Technology — Expected 2024
Relevant Coursework: Big Data Analytics, Machine Learning and Deep Learning.

Master of Computer Applications, Cochin University of Science and Technology — 2014-2017
GPA: 9.21

Bachelor of Computer Applications, Amrita Vishwa Vidyapeetham — 2011 - 2014
GPA: 9.26

EXPERIENCE

Project Engineer — Jun 2023 - Present
Centre for Development of Advanced Computing, India — *Thiruvananthapuram, KL, IN*

- Developed End-to-end big data pipeline for analytics of 112 call service.
- Created 112 Call Service - Analytics Dashboard for all States in India.
- Police Patrol Route prediction based on expected crime hotspots
- Natural Language queries to SQL for 112 Analytics report generation.

Research Scholar — Sept 2017 - Jun 2023
Cochin University of Science And Technology — *Ernakulam, KL, IN*

- Gained extensive experience in machine learning, deep learning, and big data analytics methodologies.
- Developing innovative methodologies for analyzing and interpreting complex data structures.
- Applying FCMs to address challenges in Big Data Analytics, including pattern recognition and decision-making.

SKILLS

- **DataScience/ Machine Learning/ Deep Learning**: Python, Supervised learning algorithms, Unsupervised Learning Algorithms, EDA, Feature engineering, Feature selection and extraction, Data visualization etc.
- **Python Packages and Framework:** Scikit-Learn, Tensorflow, Keras, PyTorch, PySpark, Transformers
- **Programming Environment:** Jupyter Notebook, Visual Studio Code, Spyder.
- **Big Data Tools:** Apache Hadoop, Apache Spark, Apache NiFi, Apache Hive
- **Database Technologies:** PostgreSQL, MySQL, MongoDB, SQL Lite
- **Server:** NVIDIA A100 GPU Server, Hadoop Server, HCI Server
- **Server Management Tools:** Nginx, Gunicorn, Grafana, Prometheus
- **Others:** Git, SVN, QGIS, Bash, Django

RESEARCH ACTIVITIES AND PUBLICATIONS

- Haritha, K.; Judy, M.V.; Papageorgiou, K.; Georgiannis, V.C.; Papageorgiou, E. Distributed Fuzzy Cognitive Maps for Feature Selection in Big Data Classification. Algorithms 2022, 15, 383. https://doi.org/10.3390/a15100383
- K., Haritha, Judy M. V., Konstantinos Papageorgiou, and Elpiniki Papageorgiou. 2023. "Distributed Genetic Algorithm for Community Detection in Large Graphs with a Parallel Fuzzy Cognitive Map for Focal Node Identification" Applied Sciences 13, no. 15: 8735. https://doi.org/10.3390/app13158735
- Haritha, K., Shailesh, S., Judy, M.V. et al. A novel neural network model with distributed evolutionary approach for big data classification. Sci Rep 13, 11052 (2023). https://doi.org/10.1038/s41598-023-37540-z
- Haritha, K., Judy, M.V. (2023). Performance Analysis of Distributed Algorithms for Big Data Classification. In: Venkataraman, N., Wang, L., Fernando, X., Zobaa, A.F. (eds) Big Data and Cloud Computing. ICBCC 2022. Lecture Notes in Electrical Engineering, vol 1021. Springer, Singapore. https://doi.org/10.1007/978-981-99-1051-9_13
- Haritha, K., Judy, M.V. (2021). Fuzzy Cognitive Map-Based Genetic Algorithm for Community Detection. In: Panigrahi, C.R., Pati, B., Mohapatra, P., Buyya, R., Li, KC. (eds) Progress in Advanced Computing and Intelligent Engineering. Advances in Intelligent Systems and Computing, vol 1198. Springer, Singapore. https://doi.org/10.1007/978-981-15-6584-7_39
- Developed a algorithm for Large-Scale Data Classification Using a Distributed Evolutionary Fuzzy Cognitive Map
- Developed an algorithm for Time Series Analysis using Interpretable Fuzzy Cognitive Maps

PROJECTS

112 call service - Analytics Dashboard for all States in India (CDAC)

- ***Technology:*** *Python, HTML, CSS, Django, pandas, plotly, folium, scikit-learn, TensorFlow*
- Contributed to the development of a nationwide 112 call service Analytics Dashboard using Python, Plotly, folium, HTML, CSS, Django, and pandas.
- Utilized TensorFlow, Scikit-learn, Python, and Django Rest-API to assist in building crime pattern predictive models based on geographic data.
- Collaborated with stakeholders to define analytical needs, integrate data sources, and deliver real-time visualizations for emergency service administrators.

Police Patrol Route prediction based on expected crime hotspots (CDAC)

- ***Technology:*** *Python, geopandas, shapely, scikit-learn*
- Implemented patrol route optimization using Evolutionary Algorithms, geopandas and shapely library.
- Utilized GPU for accelerated processing of route optimization and waypoint assignment based on hotspots.

End-to-end big data pipeline for analytics of 112 service (CDAC)

- ***Technology:*** *Pyspark, Apache NiFi, Kafka, Hadoop, Hive*
- Established a comprehensive big data pipeline to ingest, process, and analyze data from the 112 service.
- Leveraged PostgreSQL, Apache NiFi, Apache Kafka, Apache Spark, Hadoop HDFS, and Hive for data processing and analytics.

Natural Language queries to SQL for 112 Analytics report generation. (CDAC)

- ***Technology:*** *Python, Ollama, Llama2, Tensorflow*
- Implemented natural language query interface for SQL-based 112 Analytics report generation.
- Utilized Ollama and Llama2 LLM models for converting natural language queries into SQL queries, enabling users to interactively generate reports.

EXTRA-CURRICULAR ACTIVITIES

- Served as a resource person for multiple hands-on workshops on Big Data, Machine Learning, and Deep Learning, providing guidance and expertise to participants.

CERTIFICATIONS

- NPTEL – Introduction To Machine Learning (Online Course) - 2018
- Fundamentals of Scalable Data Science by IBM Skills Network (Coursera) - 2020
- Machine Learning With Big Data by University of California San Diego (Coursera) - 2022

ACHIEVEMENTS AND OTHER ACTIVITIES

- Qualified UGC NET JANUARY 2017 for Assistant Professor and Junior Research Fellowship eligibility.
- Second rank in Cochin University of Science and Technology for MCA (2017)
- Tenth rank in Amrita Vishwa Vidyapeetham for BCA (2014)