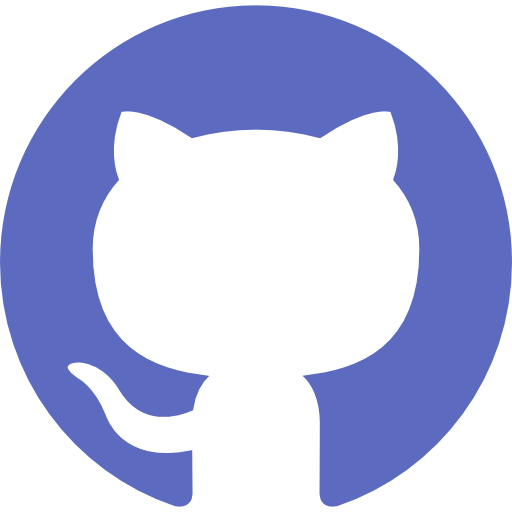
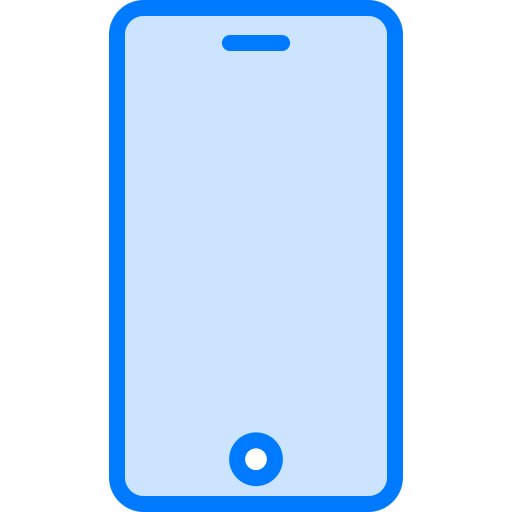
SATHISHKUMAR PARI

[LinkedIn](https://www.linkedin.com/in/sathishkumar-pari/) | +918760147068 | [sathishpari27@gmail.com](mailto:sathishpari27@gmail.com) | [GitHub](https://github.com/sathishkumarpari?tab=repositories)

61, GovindhaSamy Nagar, Kondur, Cuddalore, TamilNadu, India



**Skills**

* Python | Numpy | Pandas | PL/SQL | EDA | Data Visualization | Data Science | Statistics | Machine Learning Algorithm | Deep Learning Algorithm | NLP Algorithm | Computer vision
* Tensorflow | Keras | Scikit-Learn | ANN | CNN | RNN | A/B Testing | Time-series | Regression | Image Classification | Object Detection | Transfer Learning | NLTK | Spacy | BERT | MLOps Tools(MLFLOW,DVC) | Docker
* ODI | Oracle SQL Developer | RDBMS | ETL/ELT Pipeline | Azure | Data Bricks | PySpark | Spark SQL | Cloud Computing | CI/CD | Jira | Github
* Big Data | Data warehousing | Data Analytics | Data Mining | Predictive Modeling | English, Tamil – *All professional proficiency or above*

**Experience**

**Data Scientist II Honeywell** *Bangalore, KA, India* **12/2023 – Present**

* Developed a PoC to identify the best-fit supervised machine-learning algorithm by leveraging hyperparameter tuning techniques like Grid Search and Random Search.
* Evaluated multiple algorithms, including regression, classification, and ensemble methods, using Python libraries such as Scikit-learn and TensorFlow.
* Conducted extensive experimentation to analyze the impact of hyperparameter configurations on model accuracy and efficiency.
* Worked on Optical Character Recognition (OCR) to extract text and identify precise coordinates of text regions within images.
* Utilized OCR tools and libraries such as Tesseract and OpenCV, combined with Python, to preprocess images and enhance text detection accuracy.
* Implemented a Retrieval-Augmented Generation (RAG) system to enhance answer generation by combining Large Language Models (LLMs) with relevant information retrieval from PDFs.
* Utilized vector databases like FAISS or Pinecone for efficient document indexing and retrieval, enabling context-aware and precise responses.

**System Engineer Tata Consultancy Services** *Chennai, TN, India* **04/2021 – 12/2023**

* Unified the end-to-end data pipelines using Azure Data Factory and Azure Databricks, automating data extraction, transformation, and loading processes for enhanced efficiency.
* Transformed and optimized data warehouses in Azure Synapse Analytics, reducing query latency by 40% and enabling real-time data analytics for data-driven decision-making across the organization.
* Spearheaded the implementation of data privacy protocols and compliance standards in Azure data solutions, resulting in a 30% increase in customer trust and a 25% decrease in security incidents.
* Created and managed Azure Data Lake Storage accounts, defining data storage and access strategies for optimal usage and retrieval.
* Upgraded data processing tasks using Apache Spark and PySpark, handling large-scale data sets for analytics and reporting purposes.
* Streamlined ETL processes, reducing processing times by 30% and increasing overall data pipeline efficiency.

**Associate System Engineer Tata Consultancy Services** *Chennai, TN, India* **01/2019 - 03/2021**

* Analyzed and executed end-to-end data migration strategies using Oracle Data Integrator (ODI), ensuring a smooth data transition from source to target systems.
* Collaborated with stakeholders to define data migration requirements, mapping, and transformation rules to meet business objectives.
* Extracted data from diverse sources, including relational databases (Oracle, SQL Server), flat files, and APIs, ensuring data consistency and integrity throughout migration.
* Developed, and maintained ETL/ELT pipelines using Oracle Data Integrator (ODI) for data Integration, transformation, and loading.
* Implemented error handling mechanisms and data quality checks to validate and cleanse data, reducing errors by 60% during migration projects.

# Education

**Master of Science Chandigarh University** *Punjab, India* **09/2021 - 08/2023**

* Major: Data Science
* CGPA: 7.0/10

**Bachelor of Technology Hindustan College of Engineering and Technology** *Coimbatore, TN, India* **08/2014 - 05/2018**

* Major: Information Technology
* CGPA 6.8/10

# End To End Projects

**Deep Learning:**

* Kidney Disease Classification Deep-Learning Project with MLOps, DVC, and Deployment. [Click Here](https://github.com/sathishkumarpari/Kidney-Disease-With-MLFLOW-DVC-and-Deployment.git)
* Customer churn prediction using Artificial Neural Networks (ANNs). [Click Here](https://github.com/sathishkumarpari/Customer_churn_prediction_using_ANN.git)
* Image classification using Convolutional Neural Networks (CNNs). [Click Here](https://github.com/sathishkumarpari/Image_classification_CNN.git)

**Machine Learning:**

* Home Price Prediction: Revealing the Future of Property Values. [Click Here](https://github.com/sathishkumarpari/House_price_prediction.git)
* Exploring Factors Affecting Student Performance: A Comprehensive Analysis. [Click Here](https://github.com/sathishkumarpari/MLproject.git)

**Others**

* **Service Awards, Service & Commitment Award**: Recognition of dedicated service.
* **Contextual Master Award:** Dedication and commitment to the organization.
* **Microsoft Certified: Azure Data Scientist Associate:** Expertise in designing and implementing data science solutions on Azure.