

SUBHASRI T

Data Analyst | Data Scientist | Data Science Graduate

+91 6379942530 | ksubhasri2005@gmail.com | Chennai, Tamil Nadu (Open to Relocation)
<https://www.linkedin.com/in/subhasri-t-007b58282/> | <https://github.com/SubhasriiT>

PROFESSIONAL SUMMARY

Data Science graduate with 2 internships and 5 end-to-end projects spanning SQL analytics, machine learning, and deep learning, skilled in data cleaning, EDA, and model building for real-world business and prediction tasks. Published researcher in AI/Big Data, seeking an entry-level Data Analyst/Data Scientist role to drive data-backed decisions and build deployable models.

SKILLS

Languages & Libraries: Python, Pandas, NumPy, Scikit-learn, TensorFlow, Keras, Matplotlib, Seaborn, NLP (TF-IDF)

Databases & Platforms: SQL, MySQL, Git/GitHub, Jupyter Notebook, Streamlit

Data & Analytics: EDA, Data Cleaning & Wrangling, Feature Engineering, Statistical & Predictive Modeling, Classification, CNN, Model Evaluation & Cross-Validation, Class Imbalance Handling, Grad-CAM (Explainability)

BI & Reporting: Tableau, Power BI, Excel

Soft Skills: Analytical Thinking, Problem Solving, Communication, Teamwork, Attention to Detail

INTERNSHIP EXPERIENCE

VCodez

April 2025 - June 2025

Machine Learning Intern | [GitHub](#)

- Constructed an AI-based medicine recommendation system using NLP (TF-IDF) to automate drug type prediction from medical text across 9,700+ records, reducing manual classification effort.
- Engineered target labels and resolved class imbalance, selecting Random Forest (~64.9% accuracy) over Logistic Regression (~60.7%), achieving ~65% balanced accuracy.
- Deployed an end-to-end ML pipeline as a Streamlit web app, enabling real-time predictions across 9 drug categories with probability scores.

CognoRise InfoTech

April 2024 - May 2024

Data Analyst Intern | [GitHub](#)

- Independently executed 3 end-to-end EDA projects using Python, Pandas, and visualization libraries:
- Unemployment Analysis:** Analyzed 768 records across 28 states, identifying Tripura (28.4%) and Meghalaya (4.8%) as highest/lowest unemployment states and a sharp COVID-19-driven spike.
- Customer Personality Analysis:** Performed EDA on 2,240 customer records, uncovering wine as the top-spend category (₹675,305) and informing a 15% campaign acceptance rate.
- Salary Analysis:** Analyzed 607 data science roles, finding fully remote roles (63% of records) commanded the highest average pay, used to benchmark remote work salary premiums.

EDUCATION

UG | 2023-2026 | **B.Sc. Data Science** | M.O.P Vaishnav college for women - 7.9/10 CGPA

HSC | 2021-2023 | Muruga Dhanushkodi Girls Hr. Sec. School - 79.7%

SSLC | 2020-2021 | Muruga Dhanushkodi Girls Hr. Sec. School - 78.9%

PROJECTS

Proxy-Based Retinal Disease Severity Prediction Using CNN & MobileNetV2 | [GitHub](#)

- Built a CNN/MobileNetV2 pipeline to predict retinal disease severity by generating proxy labels from vessel density, training on the DRIVE dataset with 5-fold cross-validation, achieving 85% recall (CNN) and 100% recall (MobileNetV2) to ensure no severe cases were missed.
- Improved model interpretability via Grad-CAM visualizations across both architectures, providing a transparent, clinically trustworthy basis for predictions on limited medical imaging data.

RESEARCH EXPERIENCE

1. **Subhasri T.**, Pooja R., Sivaranjani M., & Dr. Muthakshi S. (2025). AI in Education: An Analysis of Institutional Inequities and Sustainable Adoption.

2. **Subhasri T.**, Divya M., & Dr. Aparna R. (2025). A Big Data-Driven Machine Learning Framework for Regional Temperature Anomaly Detection. ISBN: 978-93-49618-20-6

CERTIFICATIONS

- Deloitte Data Analytics Job Simulation** - Forage
- Google Cloud Computing Foundation** - NPTEL Online Certification