

## PROFESSIONAL SUMMARY

Results-oriented Data Scientist with a Master's in Computer Science and 2+ years of exposure to **Python, SQL, and Predictive Modeling**. Expert in engineering end-to-end ML pipelines (ETL, EDA) that drive business efficiency. Skilled in translating raw datasets into actionable insights using statistical analysis and visualization, reducing decision-making time by ~20%. Proficient in **Deep Learning, AWS, and CI/CD** with a proven track record of deploying models with 98% accuracy.

## TECHNICAL SKILLS

<b>Languages:</b>	Python, SQL, JavaScript, HTML5, CSS3
<b>ML &amp; AI:</b>	Supervised/Unsupervised Learning, Deep Learning, NLP, CNN, Regression, Classification
<b>Libraries:</b>	Pandas, NumPy, Scikit-Learn, TensorFlow, Keras, Matplotlib, Seaborn
<b>Cloud/Tools:</b>	AWS (EC2, S3), Git, Docker, CI/CD, Power BI, Tableau, PostgreSQL, MySQL

## PROFESSIONAL EXPERIENCE

<b>Full Stack Data Science (Apprenticeship)</b> - <i>AlmaBetter</i>	Jun 2025 – Present
<ul style="list-style-type: none"><li><b>Engineered</b> scalable Machine Learning models for 3+ real-world production use cases, boosting prediction accuracy.</li><li><b>Orchestrated</b> end-to-end ETL pipelines handling 10,000+ data points, including cleaning, feature selection, and tuning.</li><li><b>Deployed</b> a fully responsive portfolio using Astro.js &amp; Tailwind via CI/CD automation, ensuring 99.9% uptime.</li></ul>	
<b>Data Science Intern</b> - <i>The Sparks Foundation &amp; Prodigy InfoTech</i>	Mar 2024 – Jun 2024
<ul style="list-style-type: none"><li><b>Automated</b> data quality checks using Python scripts, improving dataset accuracy by <b>25%</b> and reducing manual errors.</li><li><b>Formulated</b> regression and classification models, achieving optimal performance metrics (RMSE, F1-score) on 5+ datasets.</li></ul>	
<b>Data Analyst Intern</b> - <i>Mentorness &amp; Cognorise InfoTech</i>	Mar 2024 – Jun 2024
<ul style="list-style-type: none"><li><b>Designed</b> interactive Power BI dashboards to track KPIs, visualizing 50+ key metrics for stakeholder review.</li><li><b>Executed</b> complex SQL queries to automate reporting, cutting manual data retrieval time by 30%.</li></ul>	
<b>Independent Research &amp; Upskilling</b> - <i>Self-Directed</i>	Jan 2022 – Feb 2024
<ul style="list-style-type: none"><li><b>Sharpened</b> analytical reasoning and logic through 2000+ hours of UPSC preparation and problem-solving exercises.</li><li><b>Mastered</b> foundational concepts in Python, SQL, and Statistics through 10+ completed mini-projects.</li></ul>	
<b>Associate (Operations)</b> - <i>British Airways (WNS)</i>	Dec 2019 – Dec 2021
<ul style="list-style-type: none"><li><b>Streamlined</b> workforce reporting systems, improving operational efficiency by ~15% across the department.</li><li><b>Directed</b> data-driven analysis to minimize reporting delays, saving approximately 5 hours per week.</li></ul>	

## KEY PROJECTS

<b>Kindle Recommendation System</b> (Python, SVD)	Feb 2025
<ul style="list-style-type: none"><li><b>Constructed</b> a Matrix Factorization engine (SVD) achieving a low <b>3.49 RMSE</b> score on user data.</li><li>View Code on GitHub</li></ul>	
<b>Digit Recognizer</b> (CNN, TensorFlow)	Oct 2025
<ul style="list-style-type: none"><li><b>Spearheaded</b> the development of a CNN model achieving <b>98.92%</b> accuracy on 42,000+ handwritten images.</li><li>View Code on GitHub</li></ul>	
<b>Retail Sales Optimization</b> (EDA, Pareto Principle)	Aug 2024
<ul style="list-style-type: none"><li><b>Uncovered</b> revenue patterns using segmentation and applied the 80/20 rule to identify top 20% sales contributors.</li><li>View Code on GitHub</li></ul>	

## EDUCATION

### M.Sc. Computer Science

Mumbai University | 2021

### B.Sc. Computer Science

K.M. Agrawal College | 2018

## CERTIFICATIONS

- Stanford Online: Intro to Statistics
- IBM: What is Data Science?
- CISCO: Data Analytics Essentials
- SAP: Generative AI