

Shubham Hadawle

hadawle.s@northeastern.edu | +1 8572000738 | [LinkedIn](#) | [GitHub](#) | [Android Portfolio](#) | [Portfolio Website](#)

EDUCATION

Northeastern University (Khoury College of Computer Sciences), Boston

Expected Dec 2027

Master of Science in Computer Science

Relevant Courses: Programming Design Paradigms, Database Management Systems, Algorithms, ML Ops

University of Mumbai, Mumbai

July 2020 - June 2024

Bachelor of Engineering in 'Artificial Intelligence and Data Science'

Relevant Courses: Machine Learning, Artificial Intelligence, Python, Data Structures, Object Oriented Programming – Java

TECHNICAL SKILLS

- Languages:* Python, Java, C, C++, JavaScript, R, Shell Scripting
- Mobile / Web Dev:* XML, Android Studio, HTML, CSS, Node.js, Express.js, React.js, Streamlit
- Databases:* SQL, MySQL, NoSQL, MongoDB, PostgreSQL, Firebase
- Tools & Technologies:* Git, GitHub, AWS (EC2, S3), Google Cloud Platform, PySpark, Flask, Postman, Docker, REST APIs
- Machine Learning:* Deep Learning, NLP, Computer Vision, Image Processing, Time Series Analysis, OpenCV, NLTK, LLMs, Gen AI
- Others:* Distributed Systems, Software Development Life Cycle (SDLC), Test Driven Development, MATLAB, Kubernetes

PROFESSIONAL EXPERIENCE

Motil Oswal Financial Services | *Software Engineer - Machine Learning Intern*

Sept 2025 – Nov 2025

- Developed a hybrid spatial-semantic document **clustering pipeline** to automate processing of large-scale financial research PDFs, improving retrieval quality for internal **RAG (Retrieval-Augmented Generation)** systems.
- Implemented **hierarchical clustering (HAC)** on Transformer-based text embeddings (SBERT), leveraging strong **problem solving** and **debugging skills** to generate accurate, context-preserving document clusters with 72% chunk coverage.

Axis Bank, Mumbai, India | *AI / ML Engineer Intern*

May 2025 – Aug 2025

- Worked with **GraphRAG (Graph Retrieval-Augmented-Generation)** to enhance **LLMs (Large Language Models)** by combining knowledge graphs with retrieval techniques for more accurate and context-aware outputs.
- Used GraphRAG along with Neo4j in **Text-to-SQL systems**, to help map **natural language queries to SQL** on the company's financial and investment reports by leveraging structured schema graphs, **reducing SQL querying loads**.

Bhabha Atomic Research Centre (BARC), Mumbai, India | *Software Engineer Intern - Data Science*

Nov 2023 – Jan 2024

- Used Deep Neural Networks (**LSTMs, Autoencoders, VAEs and Neural ODEs**) to solve Point Kinetic Equations that simulate the behavior of the Uranium-235 nuclear reactor.
- Utilized **Python** frameworks like **Numpy, Pandas, Matplotlib, Seaborn, Pyplot, Scikit learn, TensorFlow, Keras** and **PyTorch** (torchdyn/torchdiffeq) to automate the process of mathematical calculations thereby increasing calculation efficiency.
- Optimized cross-domain adaptation using **Adam optimizer** to refine the **loss function**, reducing the loss from **1.19520 to 0.00062**.

Stoc Guru, Mumbai, India | *Software Development Engineer Intern*

July 2021 – Sept 2021

- Java Android Developer at a FinTech startup to design & develop a social application aimed at promoting stock market investments & financial literacy, while providing real-time updates on financial news, equity, derivatives, pricing & analytics.
- Utilized **XML Scripting** Language for frontend & **Java, Python, Flask, REST APIs, testing SDKs** for backend, to test **JSON formatted data** via **Postman** to reduce latency by 2.4%.
- Implemented the backend system following the **SOLID Principles** and integrated **CI/CD pipelines** (Continuous Integration/Continuous Deployment) using Github Actions for automated builds, testing, and deployment.

PROJECTS AND PAPERS

Where is my Duck? Platformer Game

[\(Link to Mobile Game / Link to Desktop App\)](#)

- Designed and **published** a 2D platformer game for both **mobile** and **desktop (Windows/Linux/ macOS) apps** using GDevelop5, achieving **4.7/5 stars average rating** across both platforms with **1000+ downloads** on **Google Play Store**.
- Monetized the Android app** via **Google AdMob** integration to **generate revenue**, strategically deploying **Interstitial & Banner Ad** placements and conducting **A/B testing** on ad frequency, achieving a 15% increase in CTR (click-through rate).

Generating Question Paper from Handwritten Notes using T5 Transformer

- Automated the process of question paper generation from text or handwritten notes, by combining **OCR** using **Google Cloud Vision** from GCP with a **T5 Transformer** model fine-tuned for **text summarization** and **question generation**.
- Built a **Python** system to enable generation of MCQ/short-answer/True-False questions, allowing educators to upload scanned notes and download formatted question papers (PDF/Word) in < 2 minutes.

Anomaly Detection in Multi-variate Time Series Data using Facebook Prophet

[\(Research Paper\)](#)

- Developed a forecasting model for **Anomaly Detection in Time Series Data** using **Unsupervised Machine Learning** techniques and **'Facebook (Meta) Prophet'** algorithm comparing our predictions to **identify outliers**.
- Optimized model accuracy to 96.7%** through **iterative tuning** to correctly identify anomalies in online advertisement KPIs focusing on cost-per-click (CPC) and cost-per-thousand impressions (CPM).

Text Summarization of Traditional Indian Languages using Self-Attention Mechanism

[\(Research Paper\)](#)

- Headed research on **Transformer** based **abstractive text summarization** for low-resource regional Indic languages.
- Engineered a **self-attention mechanism** optimizing **positional encoding** for **morphologically rich languages**, to outperform existing **sequential model IndicBART** in **ROUGE scores** and achieving **BLEU scores** of 0.62 on a 1,000+ newspaper dataset.
- Compared results against the existing **sequential model IndicBART** achieving a **comparative ROUGE score**.