

# SUBORNO DEB BAPPON

103 Cumberland Ave S, S7N 1L6, SK

Phone: 639-916-1871 Email: subornodebbappon20@gmail.com

LinkedIn: linkedin.com/in/suborno-deb-bappon GitHub: github.com/Suborno-Deb-Bappon

## Education

### University of Saskatchewan

Sept. 2023 – Aug. 2025 (Tentative)

Master of Science in Computer Science | Grade: **89.5%**

Saskatoon, Saskatchewan

- Courses: Software Maintenance & Evolution, Empirical Software Engineering, Advanced Deep Learning, Human-Centric Software Renovation

### Chittagong University of Engineering & Technology

Feb. 2017 – Sept. 2022

Bachelor of Science in Computer Science & Engineering | CGPA: **3.76/4.00 (Merit: 3<sup>rd</sup> among 128)**

Chattogram, Bangladesh

- Courses: Data Structures & Algorithms, Database Management, AI, Machine Learning, Operating Systems

## Skills

**Programming & Query Languages:** Python, SQL, C, C++, JavaScript, HTML, CSS

**Data Analytics & Visualization:** Power BI, Tableau, Microsoft Excel, NumPy, Pandas, Matplotlib

**Machine Learning & NLP Frameworks:** scikit-learn, PyTorch, TensorFlow, LangChain

**Agentic AI Frameworks:** OpenAI Agents SDK, Crew AI

**Cloud Environment:** Microsoft Azure

**Web Development Frameworks:** Flask, Django

**Version Control Systems:** Git, GitHub

**Development Environments:** VS Code, Jupyter Lab, Cursor, Spyder

**Operating Systems:** Windows, Linux, macOS

**Practices & Methodologies:** Agile (Scrum), Data Cleaning & Transformation, Test-driven development (TDD), Debugging, Performance optimization

## Experience

### University of Saskatchewan

Sept. 2023 – Aug. 2025 (Tentative)

Graduate Research Assistant – Software Research Lab

Saskatoon, Saskatchewan

- Developed and maintained scalable data pipelines using Python (Pandas, NumPy, Matplotlib) to extract, clean, and analyze large datasets from platforms like Stack Overflow and GitHub, facilitating data-driven insights.
- Designed and evaluated transformer-based models (e.g., CodeBERT, DeBERTa) using PyTorch/TensorFlow for automated code analysis, enhancing software quality assessments.
- Implemented statistical tests (Chi-square, Pearson) to validate research hypotheses, contributing to evidence-based decision-making processes.
- Collaborated with cross-functional teams to troubleshoot data-related issues and optimize analytical workflows in a dynamic research environment.

### University of Saskatchewan

Sept. 2023 – Aug. 2025 (Tentative)

Graduate Teaching Assistant – Dept. of Computer Science

Saskatoon, Saskatchewan

- Assisted in teaching core undergraduate courses, including programming, data structures, and algorithms, supporting over 30 students per term.
- Led tutorials and guided student teams on collaborative software projects, emphasizing Agile methodologies and effective communication.
- Provided detailed feedback on assignments and facilitated office hours to clarify complex technical concepts, enhancing student comprehension.
- Supported the development of students' problem-solving skills and adherence to documentation standards, aligning with industry best practices.

### Eastern University

Feb. 2023 – Jul. 2023

Lecturer – Dept. of Computer Science & Engineering

Dhaka, Bangladesh

- Delivered lectures and lab sessions on programming, data structures, database management, and machine learning, fostering practical skills among undergraduate students.
- Mentored students on coding practices and project development, promoting analytical thinking and software engineering principles.
- Contributed to curriculum planning and departmental initiatives, aligning academic content with current technological trends.

### Extreme Solutions Ltd.

Jan. 2022 – April 2022

Software Developer Intern

Chattogram, Bangladesh

- Developed backend applications using Python and Flask, enhancing web application performance and maintainability within an Agile/Scrum framework.
- Engineered and tested machine learning pipelines for predictive analytics, improving task efficiency and supporting data-driven decision-making.
- Utilized Git for version control and collaborated with cross-functional teams to deliver scalable, testable, and production-grade software solutions.
- Applied principles of clean code, debugging, and performance optimization to ensure high code quality and maintainability.

## Selected Projects

**Power BI Data Jobs Dashboard** – Tech Stack: Power BI, DAX, Power Query, ETL, SQL, Data Modelling, UX Design

[GitHub Link]

- Designed and deployed two interactive Power BI dashboards analyzing 479K+ real data job postings, delivering actionable insights for job seekers, HR analysts, and career changers.
- Built advanced DAX measures (e.g., dynamic median salary, role-specific stats) and implemented drill-through navigation, enabling users to explore compensation, job trends, and work-from-home opportunities by role and location.
- Engineered a star schema data model with Power Query ETL to ensure scalable, clean data integration; optimized visuals with slicers, bookmarks, and KPI cards for smooth user interaction and performance.
- Created an executive-ready one-page dashboard highlighting top skills, job title salary comparisons, and global job trends—used as a decision-support tool for rapid strategic hiring and career planning.

**Neora: Multi-Agent System for Rapid Software Prototyping** – Tech Stack: Python, Crew AI, Gemini, Gradio, Docker

[GitHub Link]

- Engineered a CrewAI-based system that coordinates agents like engineering\_lead and backend\_engineer, enabling autonomous translation of plain-text software requirements into structured Python modules and functional UI prototypes.
- Designed YAML-configurable task and agent pipelines, allowing modular control over responsibilities (e.g., design generation, code synthesis) and seamless adaptation to various LLMs like Gemini 2.5 Flash for role-specific reasoning.
- Developed logic to auto-generate detailed software architecture (class/method blueprints in Markdown) followed by complete Python implementations—aligned with inferred requirements and fully unit-testable.
- Integrated Gradio-based UI rendering directly from backend logic and input schema, enabling zero-touch frontend generation and real-time interaction with generated modules.
- Built a reproducible output pipeline including session-based packaging, automatic file structuring, and ZIP export—supporting rapid MVP prototyping, testing, and Dockerized deployments.

Repolish: Automated README File Generator – Tech Stack: Python, Crew AI, OpenAI, Gemini, Gradio, CSS

[GitHub Link]

- Engineered an AI-driven documentation tool using CrewAI’s multi-agent architecture to autonomously parse repositories, extract key information, and generate professional-grade README.md files, reducing manual documentation effort by 80%.
- Developed a Gradio-based UI integrated with Gemini 2.5 Flash-powered agents, enabling users to generate, review, and deploy structured, high-quality README files — cutting onboarding time for new contributors by up to 60%.
- Designed YAML-driven task pipelines and agent configurations to support modular expansion of capabilities (e.g., markdown cleanup, PR automation), enabling rapid customization and seamless scaling across diverse repository types.
- Implemented automated GitHub pull request workflows via GitHub API, streamlining the deployment of auto-generated documentation and reducing manual update cycles from hours to minutes.

Job Market Analytics Using SQL – Tech Stack: PostgreSQL, Excel, CTEs, Joins, Subqueries, Data Aggregation

[GitHub Link]

- Designed and executed a robust SQL reporting solution to extract and analyze job market data across multiple dimensions, using advanced JOINs, CTEs, and data aggregation to deliver insights on salaries, demand trends, and optimal skill paths.
- Developed reusable, well-documented queries and scripts, ensuring high accuracy and performance while reinforcing best practices in data integrity, quality control, and enterprise-grade reporting standards.
- Translated technical data findings into clear, accessible insights through GitHub documentation—demonstrating the ability to communicate complex results effectively to both technical and non-technical stakeholders.

Coffee Shop Sales Dashboard – Tech Stack: Excel, PivotTables, Charts, Slicers, Power Query

[GitHub Link]

- Designed an interactive Excel dashboard to analyze \$698K+ in retail sales data across multiple dimensions—day, hour, product, store, and size—supporting evidence-based business decisions.
- Built dynamic visualizations and calculated KPIs (Avg Bill/Person, Orders/Person, Footfall) to uncover trends in peak hours, top-performing products, and weekday vs. weekend revenue performance.
- Streamlined reporting using PivotTables, slicers, and structured data models—demonstrating strong communication, data storytelling, and reporting automation aligned with enterprise expectations.

AUTOGENICS – Tech Stack: Python, Flask, LangChain, Gemini/GPT API, HTML/CSS, SQL

[GitHub Link]

- Developed a research-backed tool that automates the generation of context-aware, noise-free inline code comments for Stack Overflow answers using large language models (LLMs).
- Designed and evaluated an LLM-driven pipeline that processed 3M+ Q&A entries and generated comments across 400 curated Python and Java snippets—scoring high on accuracy, adequacy, conciseness, and usefulness through manual and user-based evaluations.
- Developed a browser-integrated prototype combining frontend DOM manipulation with backend Flask APIs, showcasing strong data integration, system design, and technical documentation skills.

Certifications

Data Science Foundations

- What is Data Science? (IBM)
- Data Analysis with Python (IBM)

Machine Learning and Deep Learning

- Machine Learning with Python (IBM)
- Introduction to Deep Learning & Neural Networks with Keras (IBM)
- Neural Networks & Deep Learning (DeepLearning.AI)
- Improving Deep Neural Networks: Hyperparameter Tuning, Regularization & Optimization (DeepLearning.AI)
- Structuring Machine Learning Projects (DeepLearning.AI)

Natural Language Processing (NLP)

- Natural Language Processing with Classification & Vector Spaces (DeepLearning.AI)
- Natural Language Processing with Tensorflow (DeepLearning.AI)

Problem Solving

- Problem Solving (Basic) (HackerRank)
- Problem Solving (Intermediate) (HackerRank)

Publications

- Suborno Deb Bappon**, Saikat Mondal, and Banani Roy. “AUTOGENICS: Automated Generation of Context-Aware Inline Comments for Code Snippets on Programming Q&A Sites Using LLM.” In *2024 IEEE International Conference on Source Code Analysis and Manipulation (SCAM)* (pp. 24–35). IEEE.
- Saikat Mondal, Suborno Deb Bappon, and Chanchal K. Roy. “Enhancing User Interaction in ChatGPT: Characterizing and Consolidating Multiple Prompts for Issue Resolution.” In *2024 IEEE/ACM 21st International Conference on Mining Software Repositories (MSR)* (pp. 222–226). IEEE.
- Md. Samiul Alim, Suborno Deb Bappon, Shahriar Mahmud Sabuj, Md Jayedul Islam, Muhammad Masud Tarek, Md. Shafiu Azam, and Md. Monirul Islam. “Integrating convolutional neural networks for microscopic image analysis in acute lymphoblastic leukemia classification.” *Systems and Soft Computing*, 6, 200121 (2024).
- Krishna Mridha, **Suborno Deb Bappon**, Shahriar Mahmud Sabuj, Tasnim Sarker, and Ankush Ghosh. “Explainable Machine Learning for Drug Classification.” In *2023 International Conference on Electrical and Electronics Engineering* (pp. 673–683). Springer.
- Krishna Mridha, Tasnim Sarker, Suborno Deb Bappon, and Shahriar Mahmud Sabuj. “Attention U-Net: A Deep Learning Approach for Breast Cancer Segmentation.” In *2023 International Conference on Quantum Technologies, Communications, Computing, Hardware and Embedded Systems Security (iQ-CCHES)* (pp. 1–6). IEEE.
- Suborno Deb Bappon** and Asif Iqbal. “Classification of Tourism Reviews from Bengali Texts using Multinomial Naïve Bayes.” In *2022 25th International Conference on Computer and Information Technology (ICCIT)* (pp. 270–275). IEEE.
- Suborno Deb Bappon**, Golam Sarwar Md. Mursalin, and Muhammad Ibrahim Khan. “Sentiment Analysis of Bengali Texts on Online Tech Gadget Reviews using Machine Learning.” In *2022 25th International Conference on Computer and Information Technology (ICCIT)* (pp. 324–329). IEEE.
- Suborno Deb Bappon**, Ashim Dey, Shahriar Mahmud Sabuj, and Annesha Das. “Toward a Machine Learning Approach to Predict the CO<sub>2</sub> Rating of Fuel-Consuming Vehicles in Canada.” In *2022 25th International Conference on Computer and Information Technology (ICCIT)* (pp. 384–389). IEEE.
- Golam Sarwar Md Mursalin, Suborno Deb Bappon, and Muhammad Ibrahim Khan. “A Deep Learning Approach for Recognizing Textual Emotion from Bengali-English Code-Mixed Data.” In *2022 25th International Conference on Computer and Information Technology (ICCIT)* (pp. 330–335). IEEE.

Leadership / Extracurricular

- Vice President Internal**, Computer Science Graduate Council, University of Saskatchewan (Sep. 2024 – Present)
- Mentor**, Eastern University Computing Club (May 2023 – Jul. 2023)

Honors & Awards

- Best Research Award**, CMPT 854: Human-Centric Software Renovation, University of Saskatchewan (Spring 2024)
- 75<sup>th</sup> Anniversary Recruitment Scholarship**, CGPS, University of Saskatchewan (2023 – 2025)
- Dean’s Award**, Faculty of Electrical & Computer Engineering, Chittagong University of Engineering & Technology (Sept. 2022)

References

Available upon request.