

Avi Goyal

Würzburg | Male | Indian | 13.10.2000

+49 176 87867118 | avi.goyal@stud-mail.uni-wuerzburg.de | [linkedin.com/in/avi-goyal](https://www.linkedin.com/in/avi-goyal) | github.com/GoyalAvi

PROFILE

Ich studiere derzeit im Masterstudiengang Informatik an der Universität Würzburg mit einem besonderen Fokus auf Maschinelles Lernen, Datenanalyse und Algorithmenentwicklung. Durch mein Studium und praktische Projekte konnte ich fundierte Kenntnisse in der Modelloptimierung, Computer Vision und verteilten Systemen aufbauen. Mit Begeisterung für komplexe technische Herausforderungen suche ich eine anspruchsvolle Position in einem innovativen Umfeld, in dem ich meine Kenntnisse weiter vertiefen und aktiv zum Fortschritt beitragen kann.

BERUFSERFAHRUNG

Universitätsklinikum Würzburg *Studentische Hilfskraft – Data Analyst*

Aug 2024 – Sept 2025
Würzburg

- Verarbeitete und analysierte medizinische Bilddaten (z. B. OCT) für Erkennungs- und Segmentierung Aufgaben.
- Entwickelte Machine-Learning-Modelle für die klinische Bildanalyse mit Python und PyTorch und erreichte dabei über 90 % Validierung Genauigkeit.
- Implementierte optimierte Trainings-Pipelines (GPU-beschleunigtes Training, Data Augmentierungen, Batch-Verarbeitung) zur deutlichen Leistungssteigerung der Modelle.
- Er erstellte Lehrmaterialien und unterstützte KI-bezogene Lehrveranstaltungen.
- Unterstützte die Übungen und Materialien für die Lehrveranstaltung „AI in Medicine“ und arbeitete eng unter der Betreuung von Prof. Dr. Rüdiger Pryss.
- Arbeitete mit Ärztinnen/Ärzten und Professor:innen an datengetriebener medizinischer Forschung zusammen.
- Tech Stack: Python, PyTorch, PySpark, OpenCV, Pandas, Alumentations, Matplotlib, Scikit-learn.

The Julius Maximilians University of Würzburg *Data Science Research Assistant - AutoML in Data Science and Computer Vision*

Mar 2025 – Oct 2025
Würzburg

- Entwicklung einer Produktion von automatischen Large Language Models (Deep Seek Coder 7B).
- Implementierung von fünf zentralen Erweiterung des
- Generierung und Validierung von ca. 1.900 einzigartigen neuronalen Netz Modellen und Erweiterung des LEMUR-Datensatzes um das 13-Fache.
- Erzielung von 2–12 % Leistungssteigerung gegenüber hand designten Baselines auf sieben Benchmark-Datensätzen (MNIST, CIFAR-10/100, CelebA, ImageNette usw.).
- Entwicklung des Ansatzes Few-Shot Architecture Prompting, mit +11,6 % Genauigkeit, Gewinn auf CIFAR-100 und Identifikation von $n = 3$ als optimaler Konfiguration für LLM-basierte Modellgenerierung.
- Aufbau eines submillisekunden schnellen MD5-Validierungssystems, das ca. 100 Duplikate verhinderte und 200–300 GPU-Stunden einsparen sollte.
- Analyse des Kontexts Überlauf-Problems mit 99,8 % Generierung Fehlern bei zu großen LLM-Prompts.
- Anwendung Datensätze balancierter Statistik, wodurch bis zu 15,7 % Bewertungen Verzerrung korrigiert wurde.
- Mitarbeit am Forschungsprojekt „Enhancing LLM-Based Neural Network Generation: Validation, Prompt Engineering, and Systematic Optimization“ unter der Betreuung von Prof. Dr. Radu Timofte.
- Tech Stack: Python, PyTorch, DeepSeek, LLMs, Machine Learning, Deep Learning, Generative AI, AutoML, NAS, NumPy, Pandas.

PTN Event *Softwareentwicklung*

June 2022 – Mar 2023

- Zusammenarbeit mit Teamleitern, Business Analysten und Product Ownern zur Spezifikation funktionaler und technischer Anforderungen.

- Unterstützung des Architektenteams zur Sicherung architektonischer Kohärenz und hoher Softwarequalität.
- Mitarbeit an der Erstellung technischer Entwurfsdokumente gemäß ISO/Projektstandards.
- Aktive Beteiligung an agilen Prozessen (Scrum); starke Führungsqualitäten zur Sicherstellung termingerechter Auslieferung.
- Tech Stack: PHP, Wordpress, Python, JavaScript, React, Node.js, REST APIs, SQL, Git, Docker, Agile, Jira.

The Sparks Foundation

Sept 2020 – Dec 2021

Data Science & Business Analysis Praktikum

- Umsetzung praxisnaher Projekte im Bereich überwachtes maschinelles Lernen und prädiktive Analytik.
- Entwicklung und Anwendung von Regressions- und Klassifikationsmodellen zur Lösung betriebswirtschaftlicher Probleme.
- Durchführung quantitativer Analysen zur Unterstützung datengetriebener Entscheidungsprozesse.
- Tech Stack: Python, Pandas, Scikit-learn, NumPy, Matplotlib, Seaborn, Jupyter Notebook.

EDUCATION

Julius Maximilian University of Würzburg

Oct 2023 – Present

Master's in Computer Science

CGPA: 2.4

SRM Institute of Science and Technology (Main Campus), Chennai

July 2018 – May 2022

B.Tech in Computer Science and Engineering

CGPA: 9.0/10

TECHNICAL SKILLS

Programming Languages: Go, Python, C, C++, JavaScript

Data Science & Analytics: Data Preparation, Predictive Modeling, Quality Management

Lib/Frameworks: Django, Django REST, Node JS, React JS, Flask

Databases: MySQL, MongoDB, PostgreSQL, Redis

Cloud Native Technologies: Kubernetes, Docker, Prometheus

Platform Skills: Tableau, Machine Learning, Deep Learning, NLP, Microsoft Azure

Other Skills: Analytical Thinking, Leadership, Adaptability, Time & Stress Management & Collaboration

LANGUAGE SKILLS

German (A2 Level) | English (C1 Level) | Hindi (Native Speaker) | Japanese (A1 Level)

CERTIFICATIONS

- 🌟 Google Data Studio A-Z Data Visualization (2024)
- 🌟 Data Science With Python: Data Processing and Visualization (2024)
- 🌟 HackerRank Python Certificate (2023)

PROJECTS

Petroleum Trade Network

- Developed website for the Oil and Gas Conference
- <https://ogad-conference.com/>
- <https://ptnevents.com/>
- <https://digital-twin-conference.com/>
- <https://artificial-lift-summit.com/>
- <https://container-shipping-conference.com/>
- <https://supplychain-conference.com/>

- <https://digital-transformation-conference.com/>
- Tech Stack: Django, HTML, CSS, JS, Wordpress, Redux, React Js.

Major Project – Water Quality Classification using Machine Learning

8TH SEMESTER, FINAL YEAR
PROJECT | GRADE: A+

- Datensatz mit über 7.000 Wasserproben gesammelt und verarbeitet, basierend auf chemischen Eigenschaften.
- Durchführung von EDA, Behandlung fehlender Werte, Skalierung von Merkmalen und Label-Encoding zur Modellvorbereitung.
- Entwicklung und Vergleich mehrerer Klassifikationsmodelle (z. B. Logistic Regression, Random Forest, SVM).
- Modelle anhand von Accuracy, Precision, Recall, F1-Score und Konfusionsmatrix validiert.
- Erstellung einer interaktiven Webanwendung mit Flask zur Nutzerabfrage und Ergebnisdarstellung im Frontend (HTML/CSS/JS).
- Verwendung von Jupyter Notebook für iterative Entwicklung und Visualisierung (Matplotlib, Seaborn).

Minor Project – Flood Prediction using Machine Learning Techniques

7TH SEMESTER PROJECT |
GRADE: A+

- Anwendung von überwachten Lernverfahren (Logistische Regression, Entscheidungsbaum, Random Forest, SVM) auf einen Datensatz mit über 3.000 Einträgen aus der Region Kerala zur Hochwasserrisiko Vorhersage.
- Durch Datenvorverarbeitung, Merkmalsauswahl und Modelltraining in Python mit Jupyter Notebook.
- Vergleichende Analyse der Modulleistung anhand von Accuracy, Konfusionsmatrix und weiteren Metriken.
- Praktische Erfahrung im Bereich Modelltuning, Leistungsbewertung und Interpretation realer Umweltdaten gesammelt.

PUBLICATIONS

- **Enhancing LLM-Based Neural Network Generation: Few-Shot Prompting and Efficient Validation for Automated Architecture Design**
<https://arxiv.org/abs/2512.24120>
- **NNGPT: Rethinking AutoML with Large Language Models**
<https://arxiv.org/abs/2511.20333>