

Enosh Peter Ponraj

Heidelberg, Deutschland | +49 1573 834 1757

enoshponraj@gmail.com

github.com/Enosh-P | linkedin.com/in/enosh-p

xing.com/profile/Enosh_PeterPonraj

enosh.tech



BERUFLICHES PROFIL

Software-Ingenieur und Machine-Learning-Spezialist mit über 5 Jahren Python-Erfahrung, über 2 Jahren C++-Entwicklung und einem M.Sc. in Data Science & Artificial Intelligence. Nachgewiesene Erfahrung in Voice-AI (NLU-Domain-Engineering, Spracherkennung), LLM-Integration, Backend-API-Entwicklung und CI/CD-Automatisierung. Erfahren in der Lieferung produktionsreifer Systeme bei KI-Unternehmen in Deutschland.

TECHNISCHE KENNTNISSE

Programmiersprachen Python, C++, Bash, JavaScript, TypeScript, Java, R, HTML

ML / KI PyTorch, scikit-learn, XGBoost, NLTK, spaCy, Transformers, OpenAI API, Anthropic API, Mistral, LLaMA

NLP / Sprache NLU-Domain-Engineering, Spracherkennungs-Pipelines, Tamil-NLP, Audiodatenverarbeitung, LLM-basierte Code-Generierung (CodeWeaver)

Backend / APIs FastAPI, Django, RESTful API Design, MongoDB, SQL, Prisma

Frontend / UI React, Next.js, Vue, Node.js, Tailwind CSS

DevOps / CI-CD Jenkins, GitLab CI, Docker (Grundkenntnisse), Git, GitLab, GitHub

Daten Pandas, NumPy, Feature Engineering, Regression, XGBoost

Cloud / Tools Azure (Grundkenntnisse), n8n, Linux CLI (fortgeschritten)

Sprachen Englisch (verhandlungssicher), Deutsch (B1), Tamil (Muttersprache)

BERUFSERFAHRUNG

SoundHound AI, Berlin, Deutschland

Werkstudent Software Engineer

10/2023 – 09/2025

Voice AI / C++ / CodeWeaver

- Entwicklung mehrerer produktiver NLU-Voice-AI-Domains auf Tamil in C++, um natürlichsprachliches Verstehen für Millionen regionaler Nutzer zu ermöglichen
- Einsatz von **LLM-basierter Code-Generierung** (CodeWeaver / GPT, Mistral, LLaMA) zur automatischen Erstellung von Fahrzeugsteuerungs-Grammatik-Domains
- Erhebung, Bereinigung und Verarbeitung großer Tamil-Sprachdatensätze aus verschiedenen Quellen zur Validierung von Spracherkennungsmodellen
- Entwicklung eines automatischen **Deklinations-Generators** für tamilische Nomen und Verben zur Erweiterung der NLU-Grammatikabdeckung
- Durchführung von **vergleichendem LLM-Benchmarking** (GPT-4, Mistral, LLaMA) für indische Sprachen zur Auswahl optimaler Produktionsmodelle
- Interne Schulungen zu CodeWeaver; **technische Vorstellungsgespräche** für studentische Bewerber sowie Onboarding-Seminare für neue Teammitglieder
- Enge Zusammenarbeit mit sprachübergreifenden Teams zur Standardisierung von Lokalisierungs-Workflows über 10+ Sprachprojekte hinweg

Natif AI, Saarbrücken, Deutschland

Werkstudent Backend Engineer

07/2021 – 09/2023

Python / ML / Backend FastAPI

- Erweiterung eines proprietären **Python-Test-Frameworks** für Deep-Learning-basierte Dokumentenextraktions-Pipelines mit **35 %** Leistungssteigerung
- Konzeption und Implementierung eines Aggregations-Frameworks zur Dokumentenextraktion für vielfältige Produkt- und Kundenanforderungen
- Aufbau skalierbarer **FastAPI- und MongoDB**-Backend-Dienste für produktive Dokumentenintelligenz-Pipelines
- Entwicklung maßgeschneiderter Extraktions-Use-Cases in enger Zusammenarbeit mit Fachexperten, Annotati-onsteams und Customer-Success-Abteilung
- Teilnahme an Seminaren zur Datenbankschema-Gestaltung zur Verallgemeinerung über verschiedene Dokument-typen und Anwendungsfälle hinweg

Wipro Ltd, Bengaluru, Indien

Software Engineer

07/2019 – 09/2020

Python / Bash / DevOps

- Automatisierung von Sicherheitshärtungsaufgaben zur Servermigration auf **RHEL 7**, wodurch Compliance-Aufwand erheblich reduziert wurde
- Aufbau und Pflege einer **CI/CD-Pipeline mit Jenkins** - einschließlich automatisierter Tests, Deployment-Skripte und Monitoring im Ericsson-Projekt
- Dokumentation von Automatisierungs-Workflows und Wissenstransfer-Sitzungen zur Steigerung der teamweiten Prozesseffizienz

AUSBILDUNG

M.Sc. Data Science & Artificial Intelligence

2021 – 2024

Universität des Saarlandes, Saarbrücken, Deutschland

Note: 2,4

B.E. Informatik (Computer Science & Engineering)

2015 – 2019

Anna University, Chennai, Indien

Note: 8,1 / 10,0

PERSÖNLICHE PROJEKTE

CouplesOnly — Paar-Planungs- und Tracking-Plattform

[\[View GitHub\]](#)

Tech Stack: *Next.js 14, PostgreSQL, Tailwind CSS, Python*

- Full-Stack-Lifestyle-Plattform für Paare zur gemeinsamen Planung und Alltagsorganisation
- Funktionen: Partner-Verknüpfung mit Genehmigungssystem, **Hochzeitsplaner**, Jubiläums-Tracking mit Countdown und Polaroid-Foto-Upload, gemeinsames Ausgaben- und Gesundheitstracking, To-do-Listen sowie Fotogalerie
- Architektur: Next.js API Routes (serverless) als Backend, Python-Micro-Worker zur Link-Extraktion, PostgreSQL als primäre Datenbank
- Hochzeitsplaner und Jubiläums-Bereich aktuell im **Beta-Test**

Retro-Entwicklerportfolio

[\[View GitHub\]](#)

Tech Stack: *Django, Python, JavaScript, CSS, PostgreSQL*

- Persönliche Portfolio-Website mit vollständig eigenem **CRT-/Retro-Design** der 90er: Pixel-Fonts, Scan-Line-Effekte via `repeating-linear-gradient`, dunkles Farbschema
- **Datenbankbasierte Inhalte** (kein statisches Rendering) für einfache Admin-Verwaltung
- Eigens erstellte **Mehrsprachigkeit** (Englisch, Deutsch, Tamil) ohne maschinelle Übersetzung; Sprachauswahl per Session und URL-Parameter (`?lang=en`)
- Kategoriebasierte Navigation mit **Live-Kommandozeilenterminal**, das beim Seitenaufruf automatisch relevante Skills via JavaScript aktualisiert

Tamil NMT & ASR von Grund auf

[\[View GitHub\]](#)

Tech Stack: *Python, PyTorch, Transformer, Deep Speech 2*

- **Neural Machine Translation**-Modell für Tamil-Deutsch-Übersetzung von Grund auf mit Transformer-Architektur in PyTorch — ohne vortrainierte Gewichte
- **ASR-System** für Tamil auf Basis der **Deep Speech 2**-Architektur vollständig in PyTorch implementiert
- Vertiefte Kenntnisse in CTC-Loss, Low-Resource-Sprachmodellierung und Sequence-to-Sequence-Lernen

Lieferzeitvorhersage — ML-Regression

[\[View GitHub\]](#)

Tech Stack: *Python, XGBoost, Random Forest, Pandas*

- End-to-End-ML-Pipeline zur Vorhersage von Lieferdauern aus historischen Bestelldaten
- Feature-Engineering mit temporalen, geospatialen und kontextuellen Merkmalen
- Vergleich von Linearer Regression, Random Forest und XGBoost anhand von RMSE, MAE und Kreuzvalidierung; bestes Modell zur Fehlerreduktion gegenüber der Baseline ausgewählt

AKADEMISCHE PROJEKTE

M.Sc.-Abschlussarbeit - Tamilisches Discourse Parsing (PDTB 3.0)

[\[View GitHub\]](#)

Tech Stack: *Python, PyTorch, NLP, spaCy, Transformers*

- Erstellung eines neuartigen **tamilischen Discourse-Datensatzes** von Grund auf - die erste PDTB-3.0-annotierte Ressource für Tamil mit impliziten und expliziten Diskursrelationen
- Training und Evaluation von feinabgestimmtem Transformer, LSTM-Klassifikator und SVM-Baseline für automatisches Discourse Parsing; Benchmarking nach PDTB-3.0-Metriken
- Bewältigung zentraler Low-Resource-NLP-Herausforderungen: Datenmangel, morphologische Komplexität des Tamil und sprachübergreifender Transfer

Sprecherunabhängige Sprachziffer-Erkennung (SDR)

[\[View GitHub\]](#)

Tech Stack: *PyTorch, Deep CNN, RNN-LSTM, Transformer*

- Entwicklung eines sprecherunabhängigen SDR-Systems mit Vergleich von **Deep CNN**, **RNN-LSTM** und **Transformer** auf einem Audio-Ziffern-Datensatz
- Vollständig in **PyTorch** implementiert; vertiefte Kenntnisse in **Contrastive Loss** und **Self-Supervised Learning**

Reinforcement Learning - Spielagent (Bird Hunt)

[\[View GitHub\]](#)

Tech Stack: *Python, Unity ml-agents, PPO, GAIL*

- Training eines autonomen Schieß-Agenten in einer Unity-2D-Umgebung mittels **PPO**, **GAIL** und **Behaviour Cloning** via Unity ml-agents

UniPocket - Studierenden-Webanwendung

[\[View GitHub\]](#)

Tech Stack: *ReactJS, Prisma, REST API, Node.js*

- Full-Stack-Webanwendung für Studierende zur Entdeckung verfügbarer Lehrveranstaltungen und zur Vernetzung mit Kommilitonen
- Konzeption und Implementierung von REST-Backend, Prisma-ORM-Schema und React-Frontend

ZERTIFIKATE

- R Programming for Data Science — Coursera
- Data Visualization with R — IBM