About Predii

Predii is an Enterprise AI Software Company based in Palo Alto, CA and Pune, India. Our specialized, validated AI platform has been purpose-built to extract predictive and prescriptive insights from unstructured textual, sensor, and procedural data in the automotive aftermarket and service business. Predii's patented AI engine is currently processing 2+ billion historical repair jobs monthly. Our 8+ years NLP and domain expertise enable industry leading companies such as Snap-on, Epicor, Mercedes-Benz, and Valvoline to leverage previously unused data to power predictive solutions, increase aftersales revenue, drive product innovation, and support data-driven decision-making strategies.

Predii has been recognized by Gartner, ABI Research, and the Industrial IoT Solutions World Congress for our focus in Applied AI in repair and maintenance

About this Role: Research Engineer Vision Models

As a part of your daily duties, you will be working on the Predii platform building and enhancing Vision models using state of the art Vision -Language-Models (VLMs), synthetic data generation, benchmarking, & evaluation of outputs - automatic, manual, and human-AI collaboration. There is opportunity to write papers out of the work for top-tier AI conferences. Opportunity to work with a core-AI and NLP experts with high research activity focussed on productizing and deployment at scale.

This opportunity involves developing multimodal dialog system, fine-tuning using synthetic and proprietary data sets, productizing back-end services using microservices, REST APIs, ML and NLP libraries, customized algorithms, and NoSQL datastores. You are expected to not only conduct core research but also apply to automotive ecosystem. Your role would be that of an individual contributor who will be expected to take ownership of design and develop algorithms, maintain, develop, and improve PrediiLLM models and software used in these models. You would also support in building libraries and framework that support large complex web applications. You will also use your programming skills and contribute to engineering efforts from planning to execution in development of efficient, scalable, distributed solution to real world problems. As a member of the PrediiLabs team at Predii, you will be working with cross-functional teams, customers and open-source communities. You will have the opportunity to take part in building new industry-changing services in our portfolio and participate in state-of-the-art research in LLMs

Bachelor's Degree/ Master's Degree

2-3 years experience

Competitive Salary

Pune, India

Apply now.

Basic Qualifications

- Bachelor's/Master's in Computer Science or related fields or PhD in NLP/ML/LLM (one who have submitted their thesis is eligible to apply)
- 2-3+ years of professional experience in production environment. For PhD holders, working experience in a relevant lab setting coupled with top-tier publications (A/A*) would waive the professional experience requirement

- Excellent knowledge of Computer Science fundamentals (Algorithms, Data Structures, Operating system)
- Good Programming skills (at least one of Python/Java)
- Experience in data analysis libraries such as Pandas
- Experience with LLMs fine-tuning, LLM Prompt Engineering, LLM evaluation frameworks
- Experience with Transformers architecture, Embedding models, Hugging face model publishing/inferencing
- Experience with machine learning frameworks (like Keras or PyTorch)
- Experience in machine learning, supervised and unsupervised: Forecasting, Classification, Data/Text Mining, Decision Trees, Adaptive Decision Algorithms, Random Forest, Search Algorithms, Neural Networks, Deep Learning Algorithms
- Experience in statistical learning: Predictive & Prescriptive Analytics, Regression, Time Series, Topic
- Modeling
- Basic understanding of Gradle/Maven, Git, Jira, and Confluence

Additional Qualifications

- Previous experience working with multimodal systems
- Previous exposure to vector representation of texts such as TF/IDF, Word2Vec, Glove, BERT, etc
- Knowledge of deep learning fundamentals (theory + implementation) like RNN, LSTM, CNN, Attention, seq2seq models, etc.
- Knowledge of NLP fundamentals like n-gram/BoW, NER, POS, semantics, syntax, language models, etc.
- Prior exposure to working with unstructured data and clustering algorithms
- Expertise in NLP libraries such as Spacy
- Prior exposure / Knowledge of NoSQL distributed datastores like Elastic Search, MongoDB, HDFS, etc

We're looking forward to receiving your resume and application at jobs@predii.com. For more information about Predii, please visit www.predii.com.