Sr. Machine Learning Engineer
Remote

About the position

This is a full time position that will involve retraining and implementing open source voice transcription softwares like Whisper by Open AI. You will be part of the core engineering team responsible for the design, development and deployment of Adalat AI’s algorithm trained on legal data.

Responsibilities

- Redevelop and retrain open source speech recognition transcription models, including Open AI’s Whisper
- Build interpretability tools to analyze the resulting models and understand their behavior.
- Extract insights from model results and aid in their interpretation
- Deploy models to cloud-based production environments.
- Implement other analytics solutions, such as data analysis and visualization tools to understand our client’s data and leverage it to improve our models.

About you

You are a highly enthusiastic engineer or researcher with a strong academic background. You thrive in dynamic environments and are able to take initiative. You enjoy challenges, and you have fun investigating and implementing state of the art research to solve a concrete problem. You like studying new topics as needs come up, and you’re able to communicate effectively with peers while working on highly non-linear projects.
Minimum qualifications

- Bachelor's degree in Data Science, Computer Science, Software Engineering, Mathematics, similar technical field of study or equivalent practical experience.
- At least 3 years of relevant work experience.
- Advanced Python proficiency.
- Demonstrable experience developing computer vision models in Pytorch or Tensorflow/Keras.
- Knowledge of the AWS suite and related technologies.

Preferred qualifications

- MS or PhD degree in Engineering, Data Science, Computer Science, or other technical related field; or equivalent practical experience.
- Proven ability to rapidly translate a research paper into prototypes (industry or research labs experience).
- Experience setting up large deep learning models into production environments.
- Experience designing, scaling, and optimizing machine learning infrastructure (e.g., distributed cloud computing, big data pipelines, visualization, MLOps, monitoring).

What we offer

- WFH with flexible work hours.
- Unlimited PTO.
- Contacts within the Harvard / MIT/ Oxford ecosystem.