EazyNLP: Text Mining Web App for Researchers

The EazyNLP web app is a text mining platform aimed at non-technical researchers in the social sciences who wish to take advantage of natural language processing (NLP) in the analysis of unstructured textual datasets.

Users interact with the EazyNLP web app via an intuitive and pleasing UI that they can access straight from their browser. By using point-and-click, non-technical users can seamlessly integrate EazyNLP into their research workflow by uploading datasets, conducting various NLP-based analyses, and exporting interesting findings.

Data Scientist / ML Engineer (NLP focus)

Interdisciplinary Project (or Application Project)

Your tasks:

- Contribute to the codebase of ML models (especially NLP) and data pipelines
- Implement end-to-end ML pipelines for supervised and unsupervised models, with a focus on explainability of ML/DL models (i.e., explainable AI)
- Produce intuitive, insightful and even interactive visualizations beneficial to end-users
- Thoroughly document your concepts and code in English (in-code and/or as a separate document, as appropriate)

Your profile:

- Applied programming skills in Python and familiarity with ML-relevant packages like pandas, scikit-learn, matplotlib, and NLP-focused packages (e.g. nltk, gensim, spacy)
- Ideally, some prior experience in ML (esp. NLP/text mining topics like text cleaning, vectorization, clustering), data visualization
- Previous experience in building and integrating ML models for web apps is desirable
- Ability to work independently in an agile development setup (Scrum, Kanban) and willingness to learn new approaches quickly (including version control with Git and containerization with Docker)
- Previous experience in integrating back-end code written in Python with a typical front-end setup (e.g. JavaScript, HTML, CSS, React) is appreciated
- Enough proficiency in English to clearly document concepts and code

Next steps:

- Start: asap
- Email aishwarya.kakatkar@tum.de briefly saying why you are interested in this IDP and attach your up-to-date CV and transcript of records
- Optional: send a link to your Github showcasing relevant prior projects