

Seeking NLP Research Assistant

Professor Gerard Hoberg (Finance, USC Marshall School of Business) is leading projects on the use of high-end natural language processing and AI tools. His team has worked with research assistants from CS in the past and graduates specifically noted the benefits of working on live business problems in gaining employment after graduation. We are currently in need of one computer science student who has material NLP experience to assist with NLP projects noted below. Also by applying, our team will keep your resume active should new opportunities materialize either with Professor Hoberg or by referral to other Professors doing similar work in the Marshall School of Business.

As a paid research assistant (rates range from \$20-\$25/hr), you will help our team use NLP technologies to implement robust and distributed training and inference pipelines to uncover insights from massive website text data for financial and economic analysis empowered by NLP-based transformations of content. The project aims to understand how companies are related and compete using web text and to further develop models of the full range of content present across millions of websites. You will report directly to Professor Gerard Hoberg and his team and will be supported and guided by him throughout the project.

Responsibilities:

- Understand project requirement specifications and motivation
- Read research papers in CS and identify state-of-the-art techniques to address the problem statement
- Rigorously benchmark various approaches, and document the results on our datasets using insights from the research papers and your NLP background
- Apply your NLP experience to develop and optimize training and inference pipelines that are scalable to datasets with millions of text files within the given compute resource constraints
- Write clean and readable code with detailed documentation on our GitHub repositories
- Benefit from learning about the links between CS pipelines and business applications.

Mandatory Requirements:

- Proficient in Python programming
- Prior experience with Natural Language Processing (NLP)
- Good communication and teamwork ability

Preferred Requirements:

- Experienced in machine learning frameworks, e.g. Pytorch or Tensorflow
- Experience with unsupervised learning techniques, e.g. Text Clustering or Topic Modeling
- Experience working with data processing frameworks, e.g. PySpark

How to Apply:

Please fill out [this](#) application and upload your Resume/CV.