
بازار

Information (keyword) Extraction

Description:
Information (Keyword) extraction is a text analysis technique that automatically extracts the most informative and most important words (or expressions) from a text. In CafeBazaar, we have lots of different search queries that people type into a search box in order to pull up a list of results. There are different types of queries, such as specific queries with the intent of finding a particular app or informational queries. Someone who enters informational queries, looks for the apps that answer their needs. Each of the queries in this type appears rarely but they cover lots of users' searches. There may be lots of relevant apps and we want to find the best app for the users' requirements.

Goal:
This project aims to use machine learning and natural language processing techniques to develop a model for finding and extracting informative keywords from search queries.

What you will learn:

- Developing NLP models on noisy text data
- Getting familiar with big-data/distributed frameworks, e.g., Apache Spark
- Improving your programming ability and learning to productionize your codes

Requirements:

- MS/Bs in Computer Science, Computational Linguistics, or other related fields.
- Strong programming skills in python.
- Candidates are expected to already have at least one year of hands-on experience in deep learning and natural language processing.
- Ability to demonstrate critical thinking, problem-solving, and attention to details.
- Being able to work with remote teams.
- Having experience in speech and/or natural language processing in Persian is a plus.