

Structure a Job

A job in HR domain is unstructured. It would contain title, description, location (city, state, country), posted date, id and many other fields. Assume Praveen is building a job aggregator like Monster, Naukri etc. For this he needs to scrape lots of different company websites. Let us help Praveen by building an intelligent system, which takes input a string and predicts the job entity type (like title, description etc.).

Test Cases:

1. "Hyderabad" -> LOCATION
2. "Java Developer" -> TITLE
3. "Java is a set of computer software and specifications developed by Sun Microsystems, later acquired by Oracle Corporation, that provides a system for developing application software and deploying it in a cross-platform computing environment. Java is used in a wide variety of computing platforms from embedded devices and mobile phones to enterprise servers and supercomputers. While less common, Java applets run in secure, sandboxed environments to provide many features of native applications and can be embedded in HTML pages." -> DESCRIPTION

Input:

File contains one JSON per line. Each json object contain title, description, id, posted date, category, location.

Accuracy:

- Create a test set from input data only and report accuracy.
- The system will be measured with a test set after submission.

Expected Output:

- Code
- Concise write-up/slide deck on design and implementation. (Need not be thorough).