

Targeting the rigth candidate a NLP approach

Ruddi Rodriguez Garcia

Background

Job seekers rely on social networks to search and apply for jobs. The most popular social sites are: LinkedIn, Facebook and Twitter. To increase the visibility of your job ads among potential candidates the job description has to be attractive, describing well the skills required to perform the tasks. Here I proposed a method based on Natural Language Processing and ML to improve the quality of the job ads

UnsaveApply

Seniority Level

Entry level

Industry

Marketing & Advertising,
Information Technology & Services,
Computer Software

Employment Type

Full-time

Job Functions

Engineering, Information Technology

Skills Required

- Greater than 8 years directly related experience
- Master's or PhD in Statistics, Mathematics, Computer Science
- Any certification to related area preferred
- Experience with common data science toolsets and libraries, such as pandas, keras, scikit-learn, tensorflow, numpy, matplotlib, etc.
- Experience with statistical computer languages (R, Python, etc.)
- Strong knowledge and experience using SQL language
- Experience using Azure/MS services
- Experience with C++/C# as a plus
- Excellent understanding of machine learning techniques and algorithms, such as: ML, naive Bayes, SVM, decision tree, clustering, artificial neural networks, etc.
- Strong understanding of main statistics (such as distributions, statistical testing, regression, etc.)
- Experience creating and using advanced machine learning algorithms and statistics: regression, simulation, scenario analysis, modeling, clustering, decision trees, neural networks, etc.
- Proficiency in using query languages such as SQL
- Proficiency in using query languages such as SQL
- Ability to work with internal supervision
- Minimum of 8 years' experience in area of data analysis
- Strong data analytical skills
- Creative and analytical thinker with strong problem-solving capabilities
- Background in technology or professional services preferably in one or more of the domains of: Sales, Account, Security or A/B
- Strong understanding of consulting business
- Strong structural work methods, multitasking and time management skills
- Self-driven independent work ethics that drives internal and external accountability
- May require periodic travel for workshops

Information Extraction



Web scraping: The process of data extraction from websites

- 1- Download Content of Web Pages
- 2- Parse and Extract Structured Data from it
- 3- Store data as CSV

SaveApply

Seniority Level

Associate

Industry

Marketing & Advertising,
Consumer Goods, Retail

Employment Type

Full-time

Job Functions

Other

Similar jobs

Senior Data Scientist

Tommy Hilfiger

Amsterdam, NL

Be an early applicant

1 day ago

Senior Data Scientist

PVH Corp.

Amsterdam, NL

Be an early applicant

1 day ago

Senior Data Scientist

Agoda

Amsterdam, NL

3 alumni work here

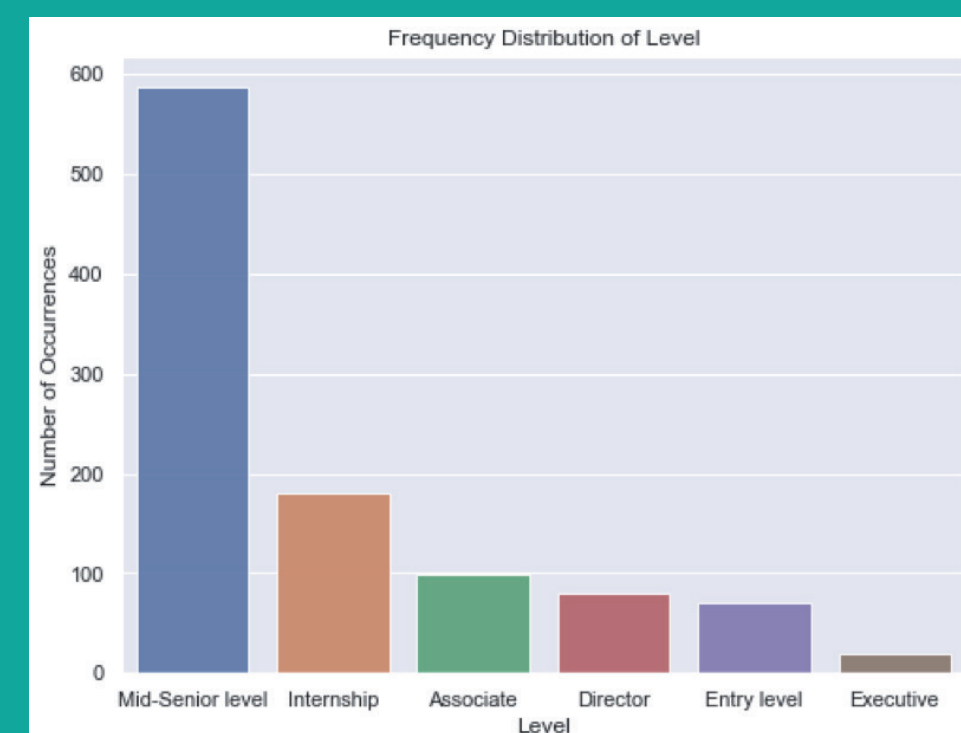
3 weeks ago

```
<!-->
<article class="jobs-description_container jobs-
description_container--condensed">
  <div class="jobs-description_content jobs-
description_content--condensed jobs-description-
content">
    <div class="jobs-box_html-content jobs-description-
content_text t-14 t-black-light t-normal" id="job-
details" tabindex="1"></div>
    <div class="jobs-description_details">
      <div id="ember1329" class="jobs-description-details
ember-view">
        <!-->
        <div class="jobs-box_group">
          <h3 class="jobs-box_sub-title js-formatted-exp-
title">Seniority Level</h3> == SE
          <p class="jobs-box_body js-formatted-exp-body">
Associate</p>
        </div>
        <div class="jobs-box_group"></div>
        <div class="jobs-box_group"></div>
        <div class="jobs-box_group"></div>
        <div class="jobs-box_group"></div>
        </div>
      </div>
    </article>
```

Data

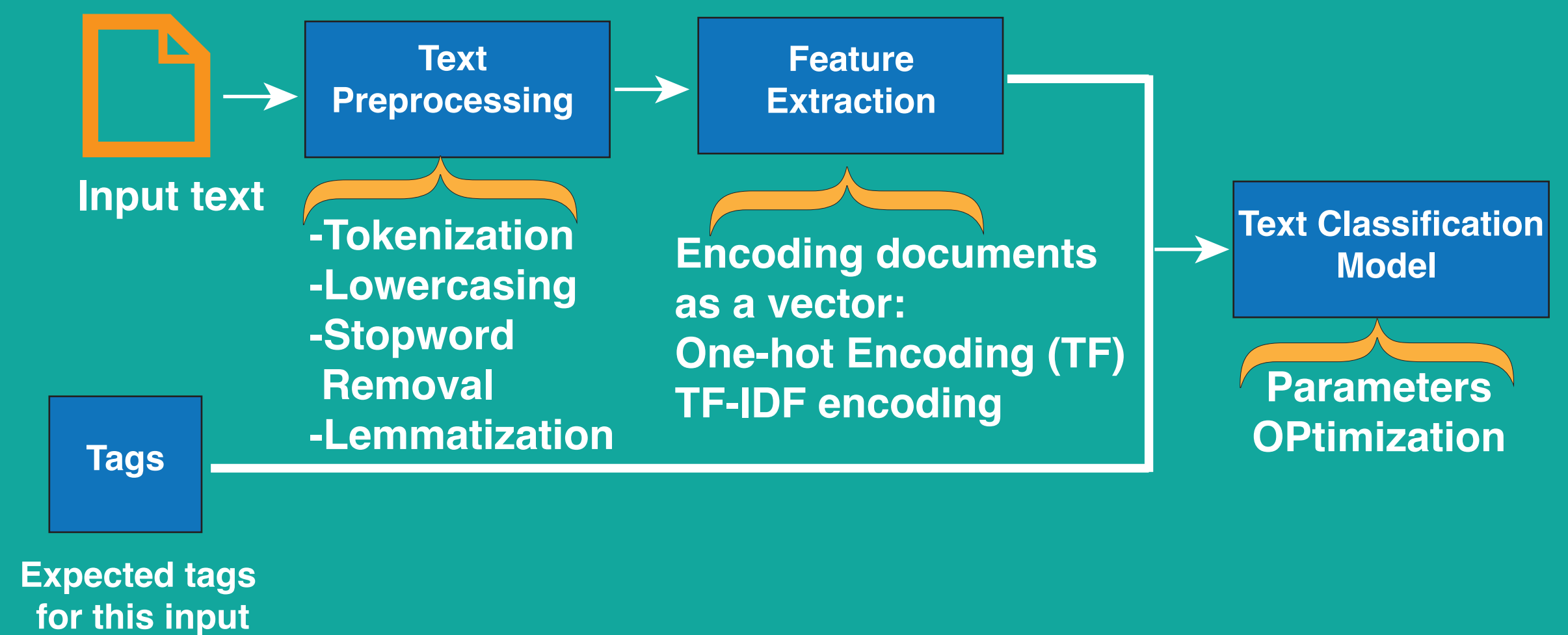
Training data: Set of 5000 jobs description posted on LinkedIn

Unnamed: 0	Level	Type_of_Job	Description
0	1	Mid-Senior level	Full-time Lead Data Engineer Data Engineer required ...
1	2	Mid-Senior level	Full-time Our client is building the future of medical d...
2	3	Not Applicable	Full-time We are looking for a Data Scientist with exper...
3	4	Not Applicable	Contract *Big data full stack engineers – who are co...
4	5	Mid-Senior level	Full-time We are seeking healthcare analytics lead to su...
5	6	Entry level	Full-time In order to be considered for this position pl...
6	7	Mid-Senior level	Full-time Senior Deep Learning Scientist Location: New...
7	8	Executive	Full-time SVP Data Science & AnalyticsHouston, TX The S...
8	9	Associate	Full-time Data Scientist is needed in Culver City, CA fo...
9	10	Mid-Senior level	Full-time Data Engineer Raleigh, NC Software Start-Up Co...



Internship
Entry Level
Associate
Mid-Senior Level
Director
Executive

Training



Prediction

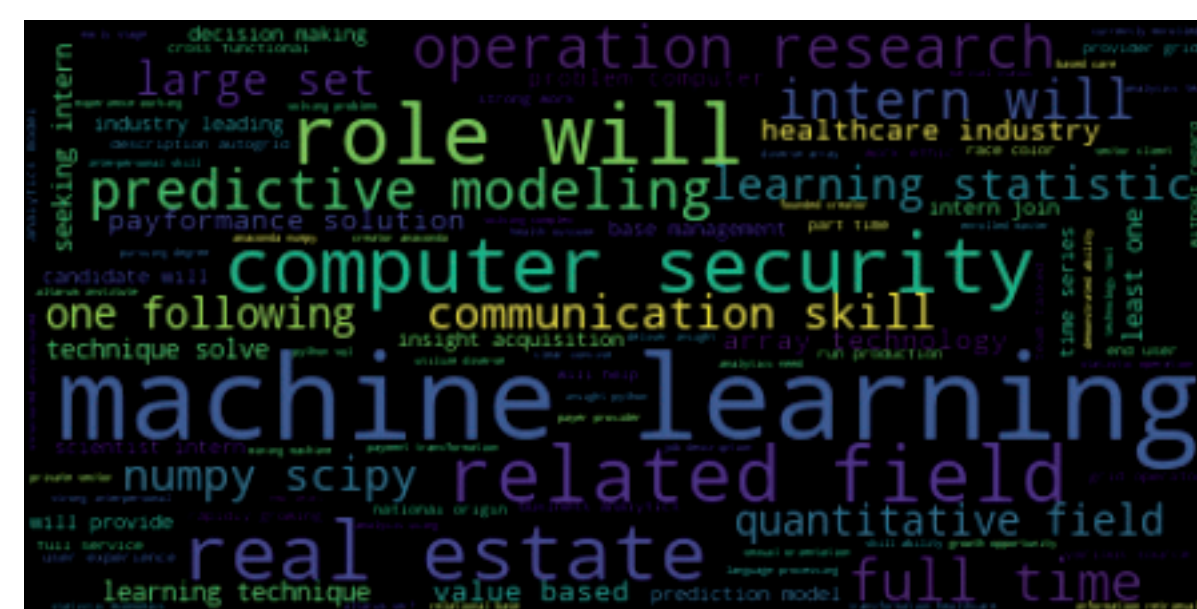


Methods

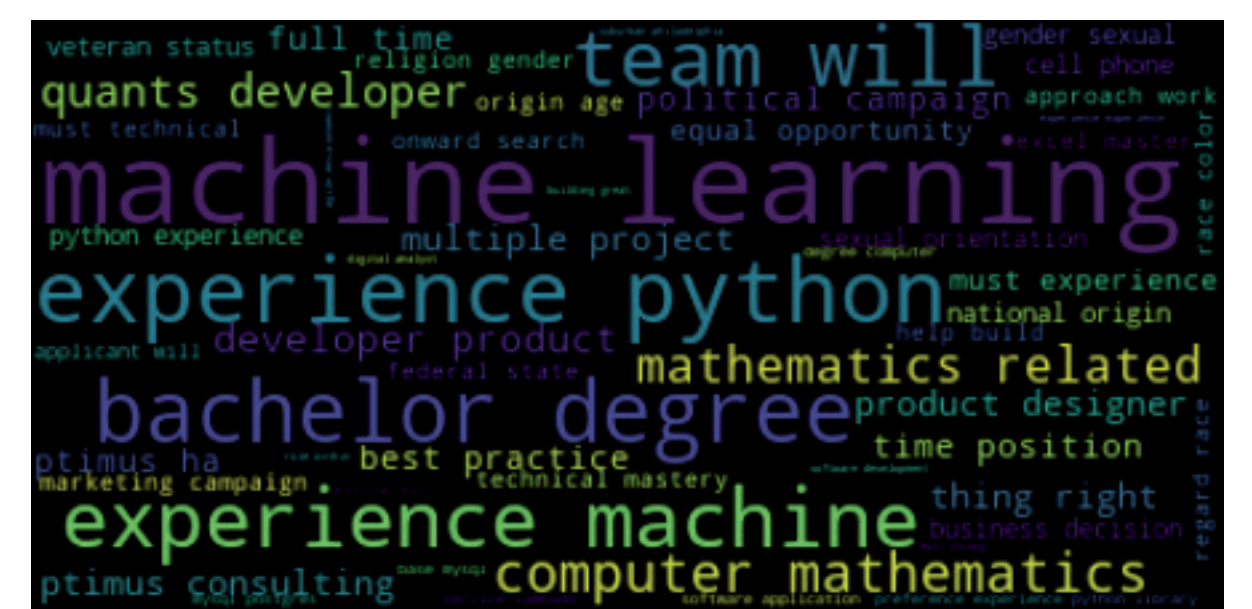
Set of python scripts: -Custom transformation Pipelines

https://github.com/RuddiRodriguez/Linkedin_analysis_NLP

Associate



Mid-Senior Level



Logistic Regression

Confusion matrix, without normalization

True label \ Predicted label	Associate	Director	Entry level	Executive	Internship	Mid-Senior level
Associate	0	0	29	0	0	54
Director	0	41	0	0	5	166
Entry level	0	0	122	0	0	251
Executive	0	0	0	9	0	13
Internship	0	0	0	0	88	63
Mid-Senior level	0	10	0	0	0	474

Naive Bayes

Confusion matrix, without normalization

True label \ Predicted label	Associate	Director	Entry level	Executive	Internship	Mid-Senior level
Associate	0	0	30	0	0	53
Director	0	84	0	0	5	123
Entry level	0	0	122	0	0	251
Executive	0	0	0	0	9	13
Internship	0	0	0	0	88	63
Mid-Senior level	0	10	0	0	0	474

Random Forest

Confusion matrix, without normalization

True label \ Predicted label	Associate	Director	Entry level	Executive	Internship	Mid-Senior level
Associate	0	0	0	0	5	78
Director	0	100	1	0	0	111
Entry level	0	0	101	0	0	272
Executive	0	0	0	9	0	13
Internship	26	9	0	0	91	25
Mid-Senior level	0	10	17	0	0	457