

NLP Resume Filter

Mitchell Seiter



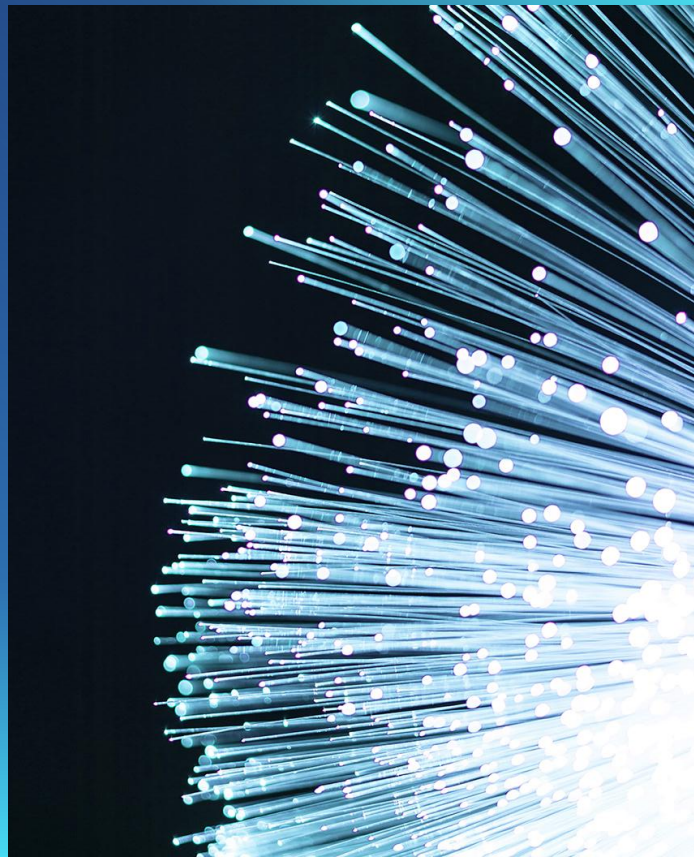


Problem Statement

It is very time consuming for recruiters and HR to sift through piles of resumes especially when many may not even be relevant to the job

Data

1000+ unlabeled resumes from
various professions



Preprocessing



01 Standardize

02 Tokenize

03 Lemmatize

04 Remove stop words (ize)

The Topics

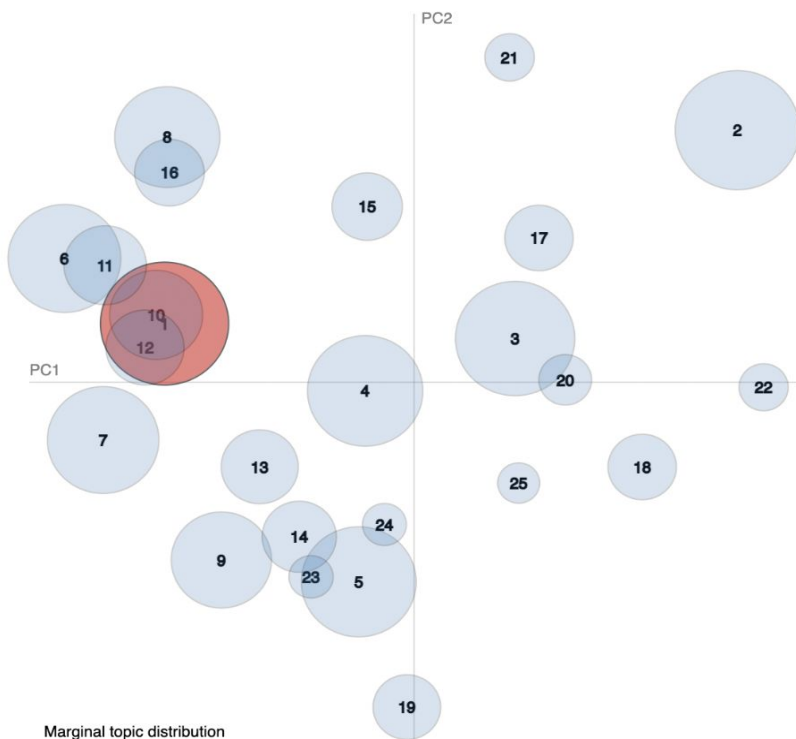
#0 - automotives?
#1 - painter/art_exhibitionist
#2 - accounting/taxes
#3 - banking/finance
#4 - film/entertainment/acting
#5 - business_administration?
#6 - graphic design
#7 - health/administration
#8 - hospitality
#9 - human_resources/payroll
#10 - teacher/educational_staff
#11 - interior_design/architecture
#12 - management/upper_executive

#13 - sales/marketing
#14 - frontend_dev/app_dev
#15 - data_analyst/DB_admin
#16 - investor
#17 - construction
#18 - social_media_specialist/digital_marketing
#19 - guidance_conselor/academic_advisor
#20 - (templates)
#21 - mechanical_electrical_engineer
#22 - fashion_designer/photographer
#23 - lawyer
#24 - doctor/nurse/hospital_worker

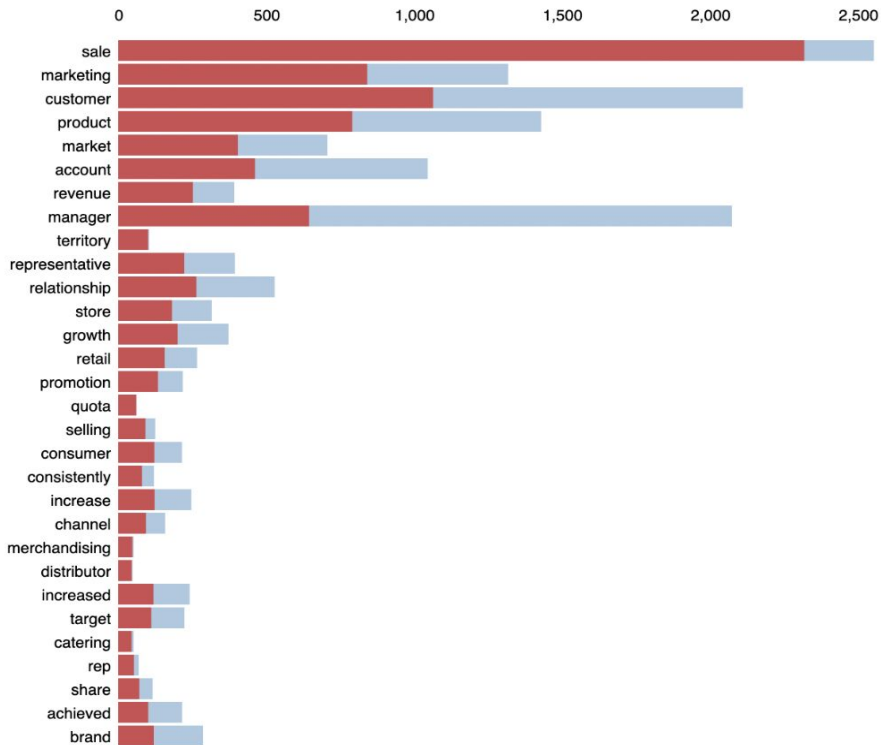


Topic Exploration

Intertopic Distance Map (via multidimensional scaling)



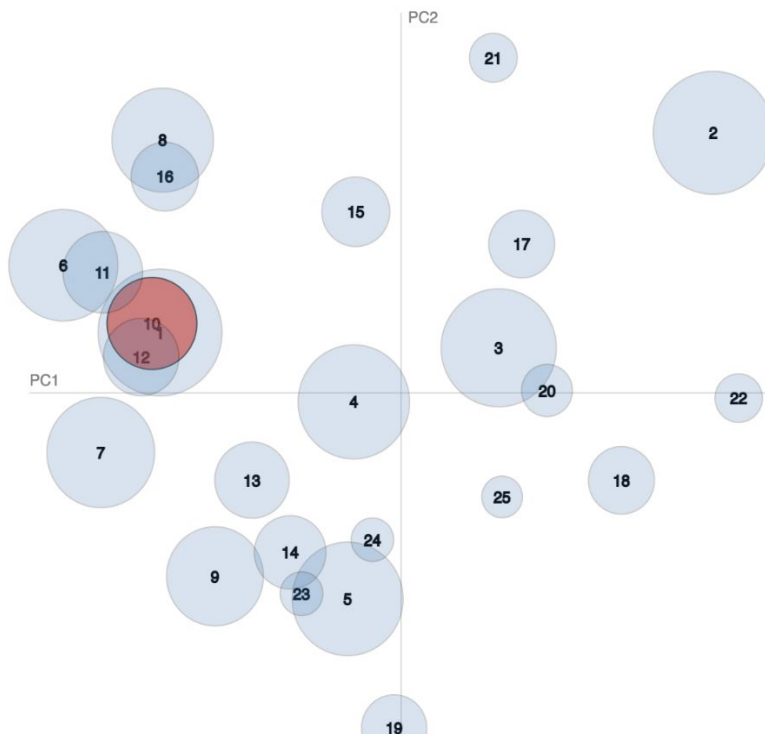
Top-30 Most Relevant Terms for Topic 1 (8.7% of tokens)



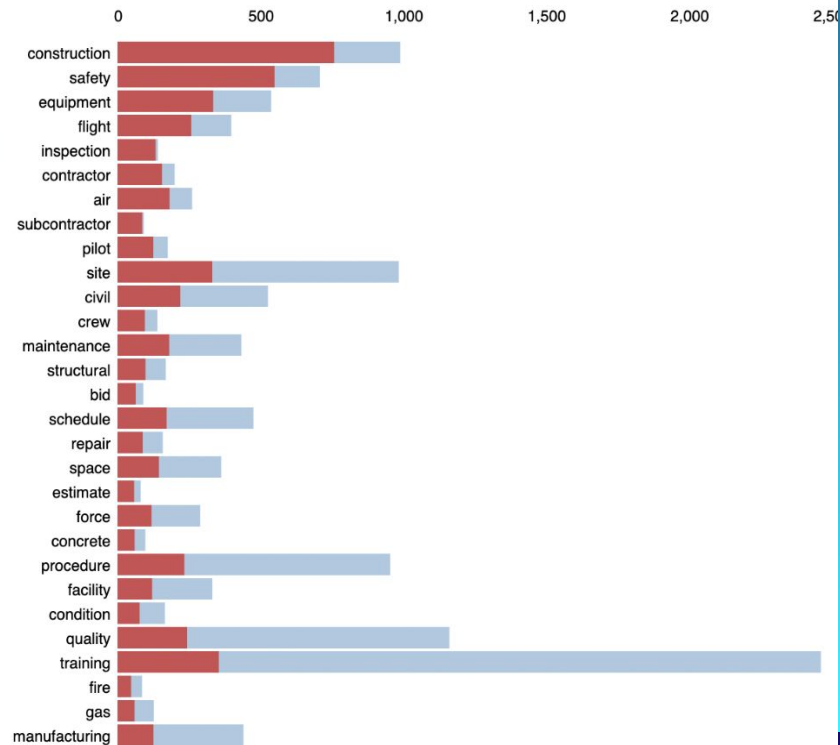
Marginal topic distribution

Topic Exploration

Intertopic Distance Map (via multidimensional scaling)

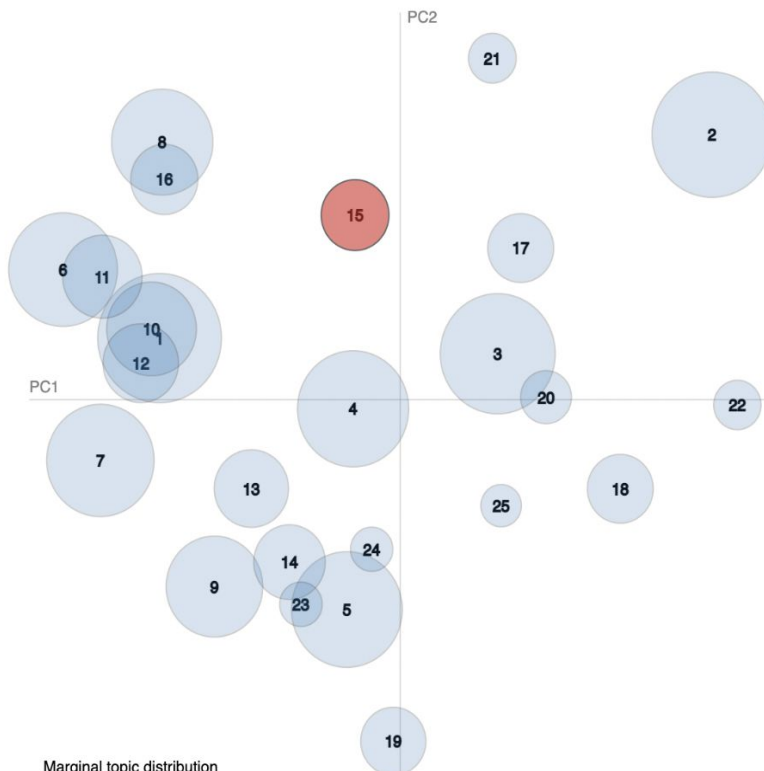


Top-30 Most Relevant Terms for Topic 10 (4.6% of tokens)



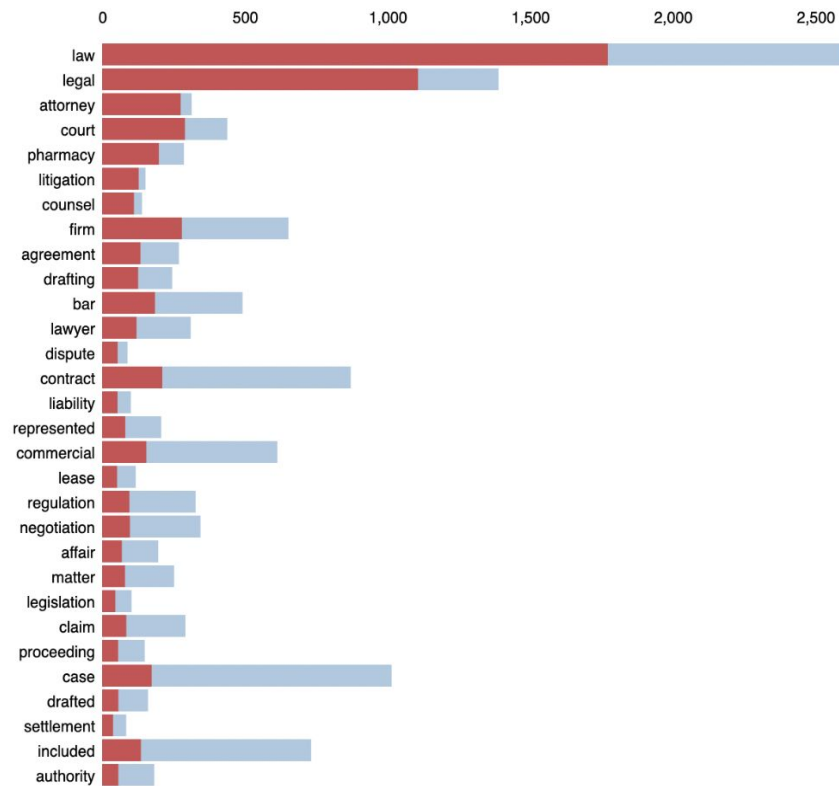
Topic Exploration

Intertopic Distance Map (via multidimensional scaling)



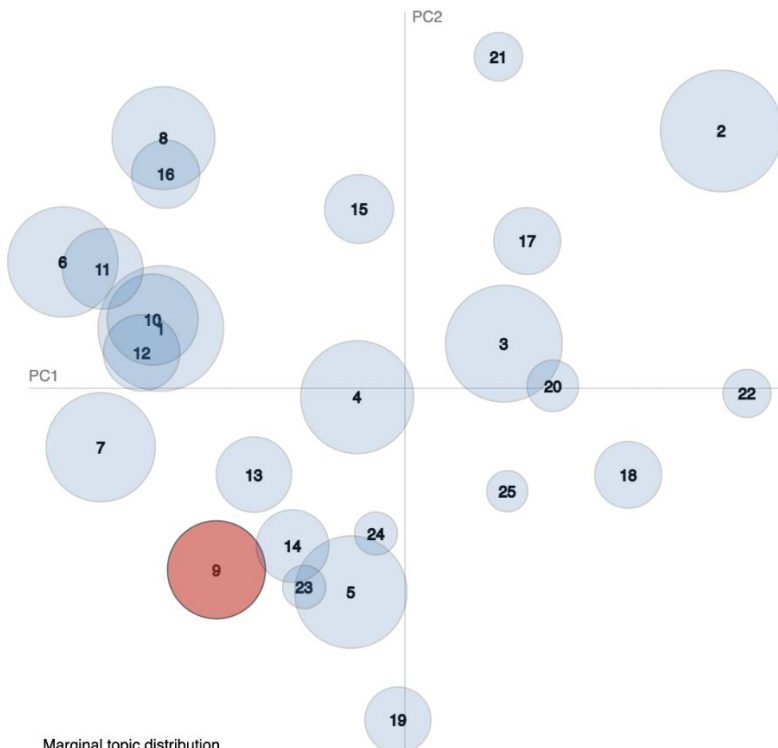
Marginal topic distribution

Top-30 Most Relevant Terms for Topic 15 (2.6% of tokens)



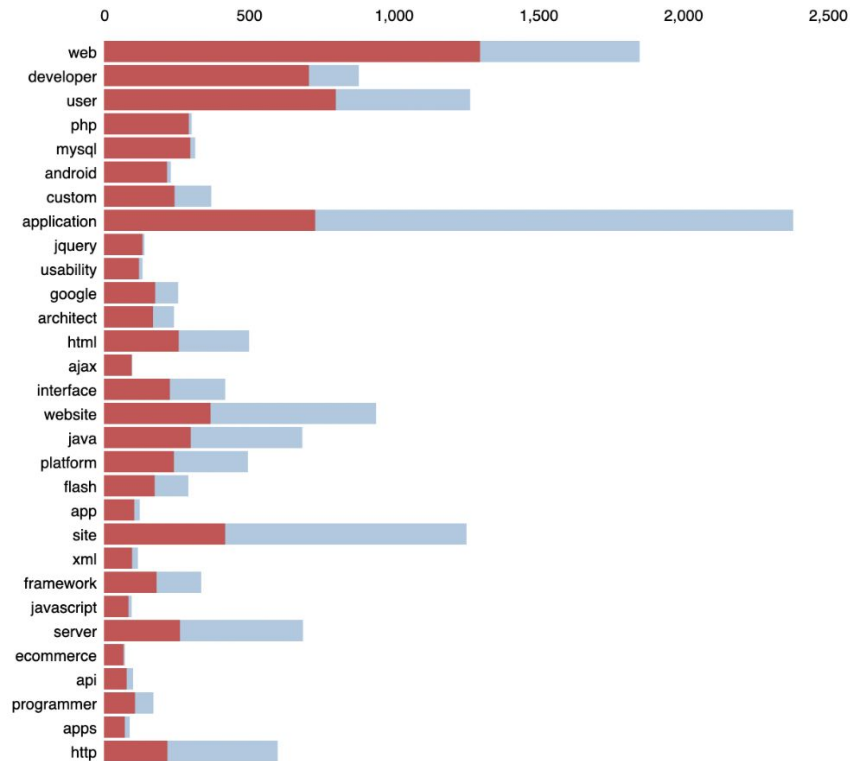
Topic Exploration

Intertopic Distance Map (via multidimensional scaling)



Marginal topic distribution

Top-30 Most Relevant Terms for Topic 9 (5.3% of tokens)



Job Posting for Software Dev

Requirements:

Bachelor's Degree in Computer Science Preferred

2-3 years industry experience at least

Knowledge of Computer Languages such as

SQL

Python

VB.net

JavaScript

Command Line/Batch Processing

C++ and C#

Goals for this position:

Learn and understand each piece of software involved in the process.

Software use involves GUI's as well as traditional line programming.

Develop a clear understanding for all the moving parts in our process in order to effectively trouble shoot problems as well as identify areas for improvement

Manage and coordinate all parties involved in changes and make necessary process changes

Assist management in the implementation of a better change management coding process for current applications.

Assist management in converting existing application to web-based applications hosted in Microsoft's Azure Platform

Assist management in the implementation support ticket system to be used by external customers as well as internal issues.



Software Engineer Resume

Andrii Chernenko Mobile Application Developer EDUCATION Period Name Dept Degree Odesa National Polytechnic University Institute Computer Systems Software Engineering Master degree Place Odesa Ukraine Present Linnaeus University School Computer Science Software Engineering Second Master degree exchange program Sweden SKILLS Programming Languages Technologies Name Skill Level Java Advanced Android SDK Advanced RESTful web service Advanced iOS Objective Intermediate SQL SQLite Intermediate Selenium Basic JMeter Basic Last Used Experience Current year Current month Current month Current month Current year month month Also some basic knowledge Python PHP Assembler Spring framework Integrated Development Environments IntelliJ IDEA Advanced Xcode Intermediate Eclipse for Java Intermediate Current year Current month month Version Control Systems Git Subversion Advanced Intermediate Current month Current year Bug tracking system Redmine Jira Advanced Intermediate Current year Current month month Operating Systems Office Software Windows Linux Mac Office Open Office Libre Office Advanced Intermediate Basic Advanced Intermediate Current year year Current month Current year year Name Russian Ukrainian English Skill Level Native Native Advanced good reading writing speaking translation skill EXPERIENCE Present Industry Position Technologies The Product Engine Inc Software Systems Telecom Junior Android Developer Java Android SDK SQLite Subversion REST JSON Responsibilities Android application development and testing Projects Android application for Jira bug tracking system us advanced REST API consumption technique and many non trivial custom component Android application for personal health management implemented integration with various data provider also done some backend development Industry Position Technologies SoGe Mobile LLC Software Systems Telecom Junior Android Developer Java Android SDK SQLite Subversion REST JSON Responsibilities Android application development and testing Projects application for restaurant chain with information about restaurant menu event and ability place order application for home appliance chain information about merchandise discount nearest store etc application for sport newspaper featuring app purchase application for travel magazine with article event and other information about travelling includes photo bank Advanced use Google Maps API application for online shop aggregator providing information about merchandise ability buy save item wish list application for making and sharing note text audio video Implemented some advanced medium feature ability split merge audio clip used complex animation custom component application for English learner includes test Industry Position Technologies Odesa Ukraine The Product Engine Inc Software Systems Telecom Junior Engineer Java TestNG Selenium Ant SQL XML XPath Subversion Responsibilities web application functional and performance testing development automated functional test suite for web application test documentation maintenance test plan test case test report Projects online advertising application performed functional and performance testing developed automation test suite maintained test documentation Odesa Ukraine Odesa Ukraine INFORMATION Courses taken Software Quality Linnaeus University Sweden Android for Java Programmers Linnaeus University Sweden Compiler Construction Linnaeus University Sweden Software Testing The Product Engine Inc Odesa hour Programming Java ONPU hour Software Design ONPU hour Stackoverflow com account http stackoverflow com user deville mostly active tag Github account http github com deville Habrahabr account http habrahabr user deville ABOUT honest careful perfectionist hard working fast learner able think logically good problem solving skill able work team good communication skill'

Lawyer Resume

'phd attorney law curriculum vitae sondra gmail pub sondra faccio current present associate attorney corporate department partner padua lecturer commercial law course post graduate student education phd corporate law economics doctor europaeus law school indirect expropriation international investment law state regulatory power investor protection master diplomacy institute international political study ispi milan master law magna cum laude law school italy degree law science law school professional traineeship trainee lawyer studio legale curtarello monselice padua civil law contract law trainee lawyer internship law lawyer notary law contract law intern political relation general consulate organising meeting attending apec economic leader meeting award ernst mach grant worldwide oead phd scholarship reministry education scholarship visiting student researcher experience postdoctoral researcher international investment law law school head cross border business management italy foreign network unicredit bank veneto visiting phd student universit prof milo vec prof reinisch visiting phd student law school new york course international finance monetary policy prof goodspeed international investment law prof prof law prof sauvant research assistant vale law school new york visiting student course attended international commercial law politics international law visiting student universit lyon lyon course attended communautaire droit international public droit compar publication article report newsletter energy reform source international investment dispute leiden application principle proportionality ass compensation reflection arising case law practice international court tribunal tra art delle preleggi loom diritto internazionale privato processuale acteurs non le etats souverains dans i financed realized partnership head cross border business management italy foreign network unicredit bank italy diffused network exclusive economic zone partner publication http fileadmin newsletter pdf energy reform challenge strategy foreign investor partner publication forthcoming http language mother tongue proficient knowledge toefl certificate proficient knowledge dalf certificate[sydney australia, sydney australia, italy cooperint fellowship international research bureau, verona italy worldwide study program scholarship international exchange student office university, european union, wien austria department international law international relation, columbia university, johnson international investment arbitration, bekker policy international, columbia center sustainable international investment columbia, united state, macquarie university, sydney australia, journal international law forthcoming, italia rivista, droit international, nella giurisprudenza, investimenti internazionali bocconi, quadro internazionale research project, european social fund]

Future Work

Step 01



More
preprocessing

Step 02



Add more
resumes to
dataset

Step 03



Give job
description
words more
weight

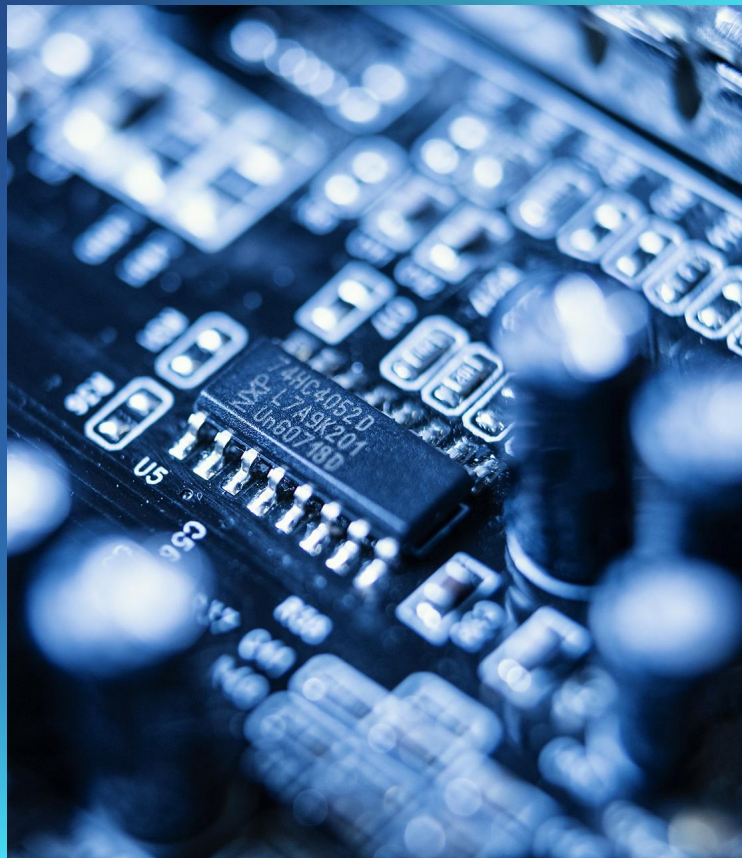
Step 04



Validate on
supervised
learning

Conclusion

- Algorithm can be used to pull resumes closely related job description if job description has good matching keywords. May start to pull irrelevant resumes if topic words are too universal.



Thanks

