



STEVENS
INSTITUTE *of* TECHNOLOGY
THE INNOVATION UNIVERSITY®

Indeed Resume Analysis

BIA-660 Web Analytics

Manank Valand
Mehul Mistry
Ayushi Chaturvedi
Shefali Chhabria





Contents of the project





Motivation & Objective

❖ For Students:

It will help the student to analyze their resume and further help them to predict their chances of getting the desired job.

❖ For Companies:

- It's one step further of Application Tracker tool which gives you visual representation of all the applications.
- Simply uploading of resume will help the companies to sort and analyze them according to the requirements.

Objective: A tool for analyzing and predicting the resume



Introduction

- ❖ This project aims to develop a tool for accurately analyzing resumes based on work experience, education details, work auth, skills etc.
- ❖ Dataset: We have scraped almost 3000 resumes from [indeed.com](https://www.indeed.com) for 3 major job titles. i.e.
 - ❖ Data Analyst
 - ❖ Senior Software Engineer
 - ❖ Vice President
- ❖ Models Tested: Deep Learning with CNN and Naïve Bayes Classification.



Scraping (1/2)

Data Scientist

CURRENT JOB TITLE

Data Scientist/Mathematician/Statistician

San Diego, CA

OBJECTIVE/ SUMMARY

To further my career in mathematics by obtaining an analytic position in data, software development, and/or strategy.
Authorized to work in United States for any employer

Work Experience

Data Scientist

PREVIOUS JOB TITLES

Sleepscore Labs - Carlsbad, CA
February 2017 to Present

Marketing Data Scientist

Thermo Fisher Scientific - Carlsbad, CA
July 2016 to December 2016

**COMPANIES WORKED/
HOW MANY YEARS?**

- Use SQL Developer and Report Studio to upload, update, and query large databases for marketing data.
- Use R to create models using machine learning which include linear regression, logistic regression, decision trees, recommender systems, clustering, and linear discriminant analysis.
- Cross-sell models are created using customer features and purchasing history.
- Web behavior models are created using page names, URLs, and impressions.
- Explain analysis and models to others in marketing and sales.

Data Scientist

Univfy - Los Altos, CA
March 2012 to June 2015

JOB DESCRIPTION

- Searched, organized, and analyzed clinical data regarding in vitro fertilization.
- Used statistical analysis to determine data behavior.
- Used machine learning on the data to create predictive models regarding IVF outcomes such as live



Scraping (2/2)

Education

Stanford Center for Professional Development in Data Science
Stanford University Stanford, CA
January 2014 to January 2015
Master of Science in Applied Mathematics
Santa Clara University Santa Clara, CA
December 2011
Bachelor of Science in Mathematical Sciences
Stevens Institute of Technology Hoboken, NJ
May 2005

EDUCATION DETAILS

Additional Information

SKILLS: Hardware: PC, Mac
Software: Microsoft Word, Excel, Eclipse, RStudio, SQL Developer, IBM Cognos Report Studio, Com
ergent Visual Modeler
Programming Languages: C++, C, R, Java, Matlab, SQL, Python
Operating Systems: Windows (any version), DOS, UNIX, OS X Lion

SKILLS

Dataset



['Stevens Institute of Technology - Hoboken, NJ', 'Hangzhou Testing Alliance Technology - Hangzhou, CN']											
TYPE	TITLE	SUMMARY	WORK_AUTH	SKILLS	TITLES	COMPANIES	DURATION	DESCRIPTION	PLACES	EDUCATION	
1	Data Scientist	Data science	Authorized to	SQL (5	yea	['Data Scientis	['BarclayCardUS - V	['May 2016 to F	['Created SQL t	['Fairleigh Dickinson	['Masters in Computer Science'
2	1 Data Scientist	NA	Authorized to	Python (4	['Data Scientis	['Foreal Spectrum -	['August 2017 t	['Designed multi	['New York Unive	['Master of Science in Data Scie	
3	1 Data Scientist	NA	Authorized to	Python (4	['Data Scientis	['Foreal Spectrum -	['August 2017 t	['Designed multi	['New York Unive	['Master of Science in Data Scie	
4	1 Data Scientist	Around 8+ yea	Authorized to	NA	['Data Scientis	['State Street - Bos	['February 2017	['Description: St	['NA']	['Bachelor's"]	
5	1 Research Assi	NA	Sponsorship r	DATA MIN	['Research As	['Stevens Institute d	['September 20	['Contribute ti	['Stevens Institute	['Master of Science in Compute	
6	1 Lead Data Sci	Proven skills an	Authorized to	PREDICTIV	['Lead Data Sc	['Comcast Cognitiv	['2016 to Prese	['Developing cog	['Dalhousie Unive	['Ph.D. in CS, EE']	
7	1 Data Scientist	Professional q	NA	NA	['Data Scientis	['Mass Mutual - Spr	['February 2017	['Description:\ N	['Informatica Pow	['Bachelor of Computer Science	
8	1 Lead Data Sci	Accomplished t	NA	Data Scien	['Lead Data Sc	['MasterCard - Bell	['April 2017 to f	['Leading a team	['University of Ari	['Bachelor's in Science in Mathe	
9	1 Principal Data	NA		SQL (7 yea	['Principal Dat	['Pilot Flying J - Kno	['January 2014	['Created severa	['American Militar	['Masters Intelligence Studies in	
10	1 Lead Data Sci	Experienced	WORK AUTH	MatLab (1	['Lead Data Sc	['Verisk Analytics',	['December 20	['Lead efforts to	['McGill University	['Ph.D. in Computational/Math	
11	1 Data Scientist	Over 8+ ye		SQL (7 yea	['Data Scientis	['CGI Group Inc - N	['February 2017	['Company Desc	['NA']	['Bachelor's"]	
12	1 Data Scientist	Highly motivate	NA	SQL (4 yea	['Data Scientis	['ESPN - Bristol, CT	['January 2016	['Understand p	['Columbia Uni	EDUCATION LEVEL	
13	1 Data Scie			NA	['Data Scientis	['Nationwide Insur	['November 20	['Project Overvie	['SOUTH DAKOTA	Master of Science in Data Scie	
14	1 Data Scie	SUMMARY/OBJECTIVES		DATABASE	['Data Scientis	['Avera Health - Sio	['May 2017 to J	['Analyzed and	['SOUTH DAKOTA	Master of Science in Data Scie	
15	1 Data Scie			NA	['Data Scientis	['Active Health Mar	['September 20	['Alpharet	['Bachelors of Technology in Te		
16	1 Data Scie	Multiple years	Sponsorship r	Apache (Le	['Data Scientis	['AT&T - Atlanta, G	['August 2015 t	['Objective	['Masters in Computer Science		
17	1 Data Scie	Above 8+ year	NA	SQL (6 yea	['Data Scientis	['AT&T - Dallas, TX	['July 2017 to P	['Description:\ A	['NA']	['Bachelor's"]	
18	1 Data Scie	Over 8+years	Authoriz	SKILLS	['Machine Lea	['Citibank - Irving, T	['February 2017	['Description: Ci	['NA']	['Bachelor's"]	
19	1 Data Scie	Having 8+ ye	Authorized to	NA	['Data scientis	['life Scan Inc', 'Car	['June		['ca Pow	['Bachelor of Computer Science	
20	1 Data Analyst	A highly motive	NA	HTML	['Data Scientis	['Verizon - Townshi	['Janua		['on Univ	['MS in Information Systems an	
21	1 Data Scientist			CVS - Woonsocke	['Data Scientis	['CVS - Woonsocke	['August 2017	['Description:CV	['NA']	['Bachelor's"]	
22	1 Data Scientist			Walgreens -	['Data Scientis	['Walgreens -		['Description: Th	['NA']	['Bachelor's"]	
23	1 Data Scientist			TriNet Group	['Data Scientis	['TriNet Group		['Description:Tri	['Informatica Pow	['Bachelor of Computer Science	
24	1 Data Scientist			Arine - San F	['Data Scientis	['Arine - San F		['startup in Calif	['Indiana Universi	['Master of Science in Data Scie	
25	1 Data Scientist			NA	['Data Scientis	['NA' 'NA' 'N	['MUSIC PLAYER D	['February 2018	['Preprocessed	['ECOLE CENTRAL	['Master's', 'Master of Science
26	1 Data Scientist			+ (3 yea	['Data Scientis	['NA' 'NA' 'N	['MUSIC PLAYER D	['February 2018	['Preprocessed	['ECOLE CENTRAL	['Master's', 'Master of Science
27	1 Data Scientist			Authorized to	DATA MOL	['Data scientis	['Ecom Consulting	['June 2013 to f	['Background\ Vi	['NA']	['Master's"]
28	1 Data Scientist			Authorized to	DATA MOL	['Data scientis	['Ecom Consulting	['June 2013 to f	['Background\ Vi	['NA']	['Master's"]

1. DATA SCIENTIST
2. SENIOR SOFTWARE ENGINEER
3. VICE PRESIDENT

PREV.
JOB TITLES
ARRAY

COMPANIES WORKED

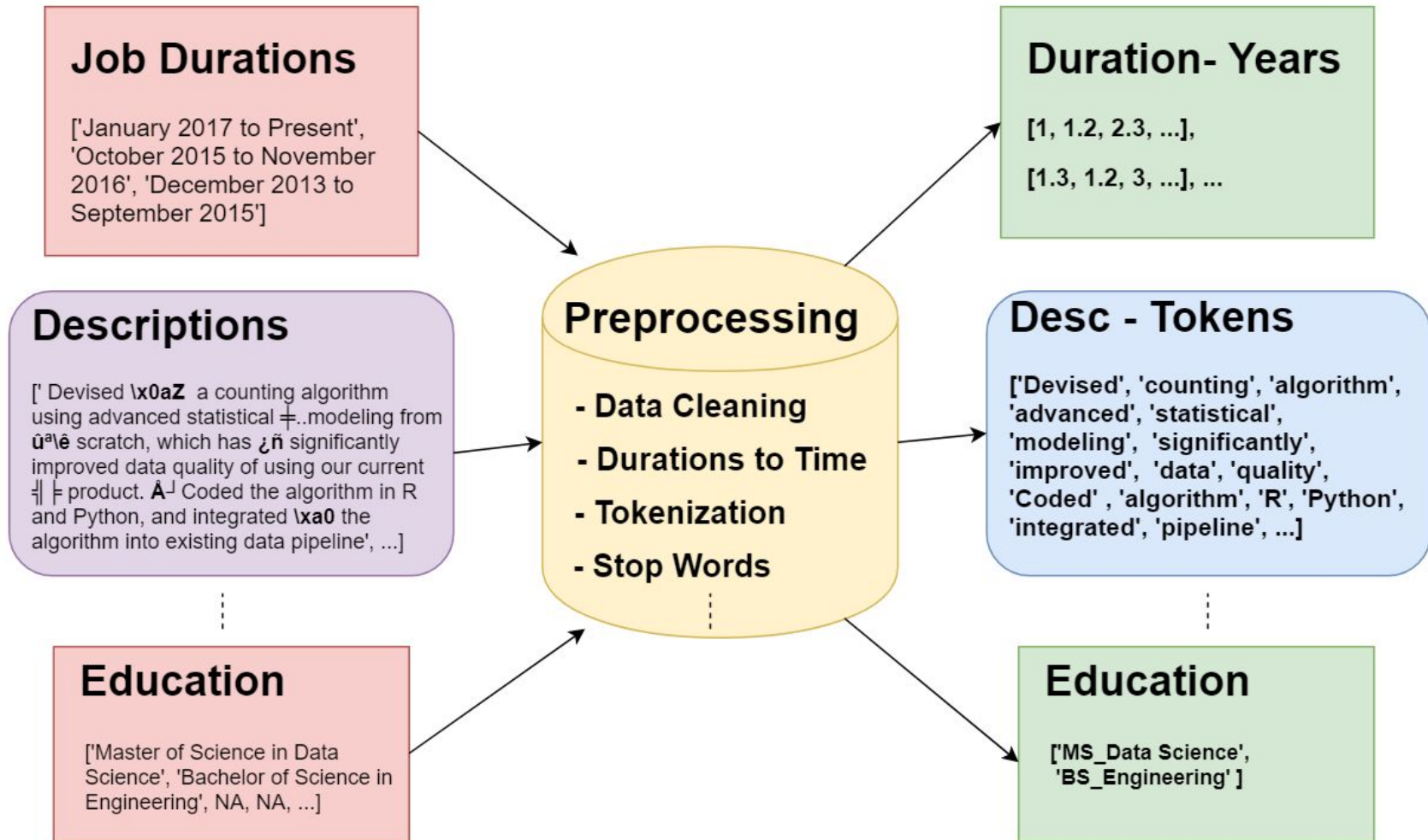
JOB DESCRIPTION

DURATION
WORKED

UNIVERSITIES

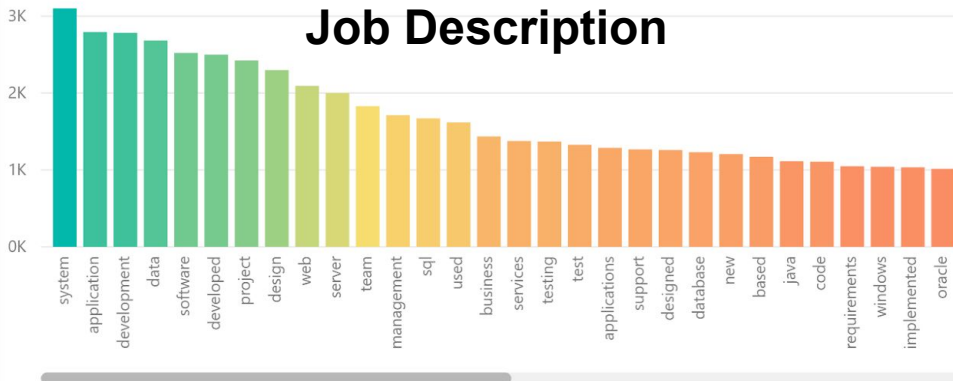
EDUCATION
LEVEL

Preprocessing



Dataset Analysis (Sr Software Engineer)

desvalue and desvalue by des

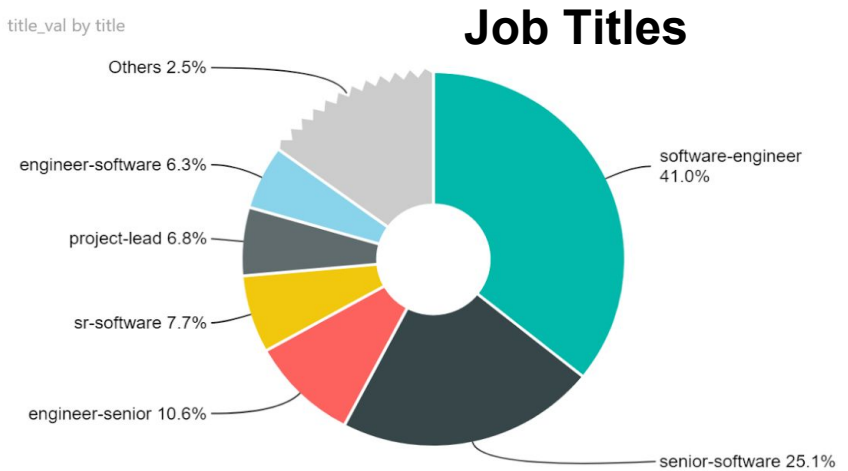


skillval by skill

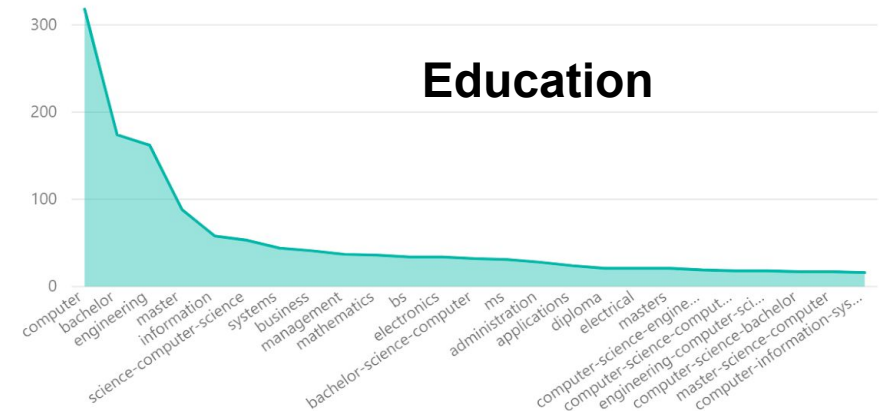


Skills

```
title_val by title
```



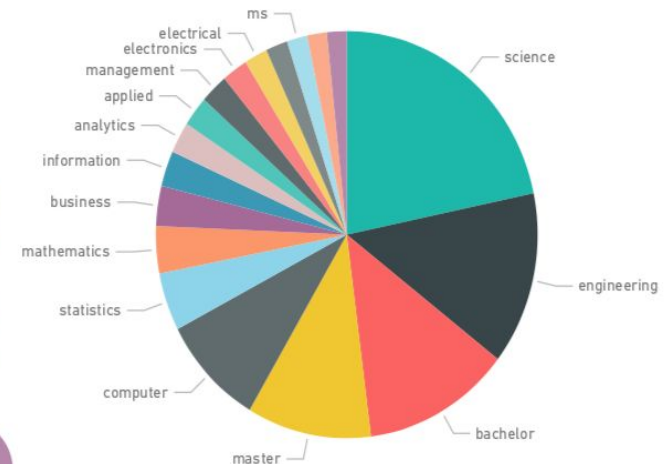
edu_value by edu



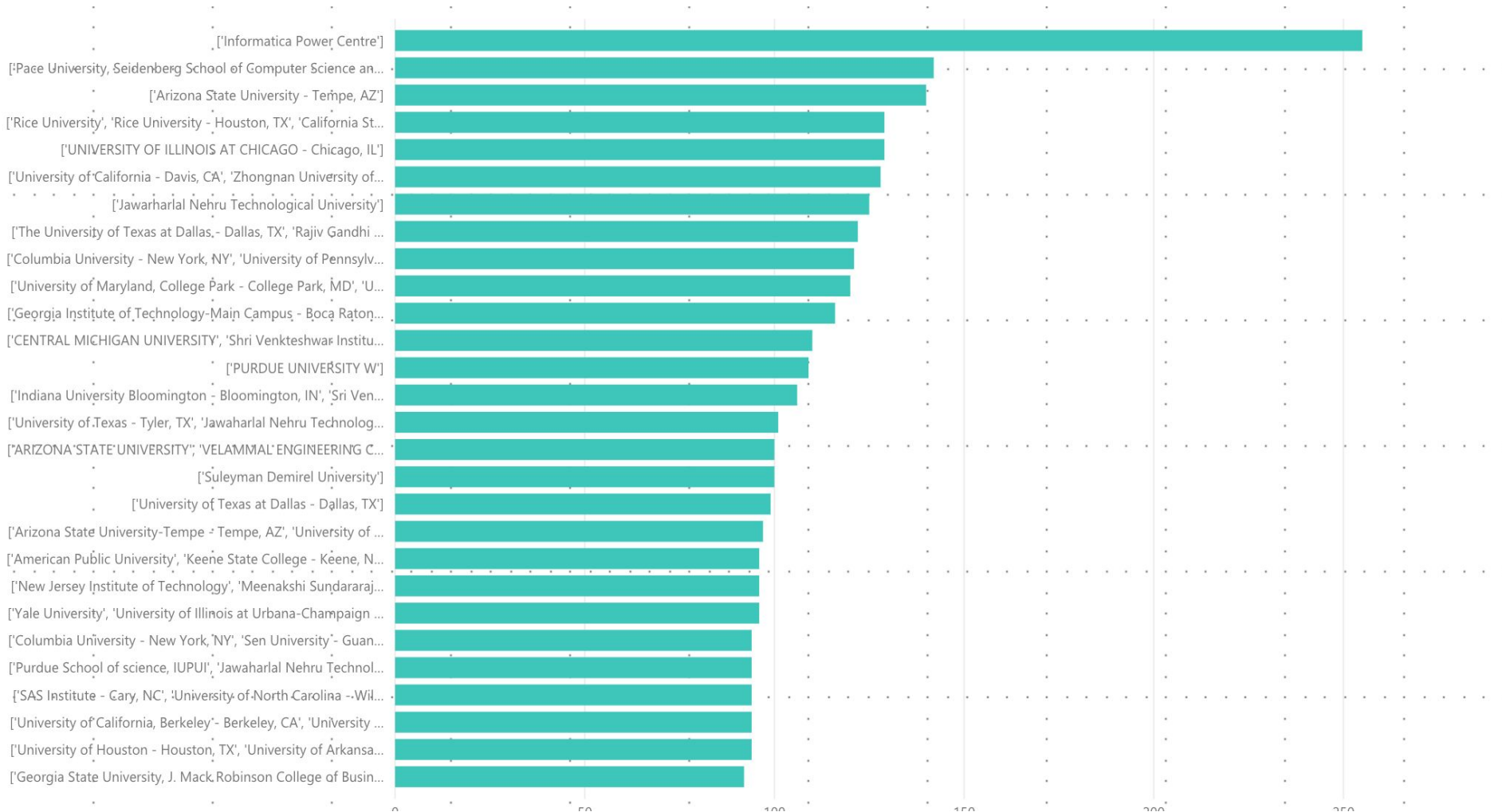
Dataset Analysis (Data Analyst)

Column6 by Column5

Column10 and Column10 by Column9

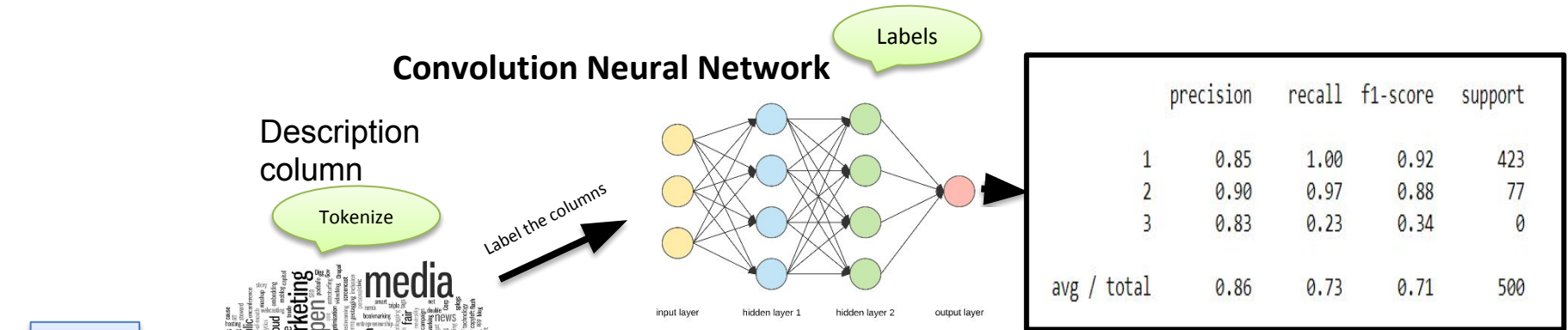


Dataset Analysis (Data Analyst)

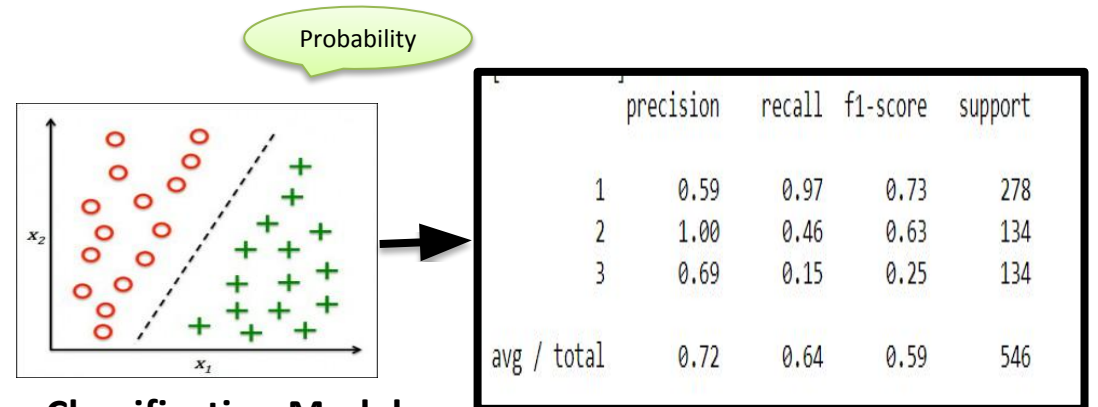


Classification Model

Convolution Neural Network



Structured + Unstructured Data



Naïve Bayes Classification Model



Experiment Results

- Performance Evaluation

- ❖ Naïve Bayes

The performance of classifying label:

	precision	recall	f1-score	support
1	0.59	0.97	0.73	278
2	1.00	0.46	0.63	134
3	0.69	0.15	0.25	134
avg / total	0.72	0.64	0.59	546

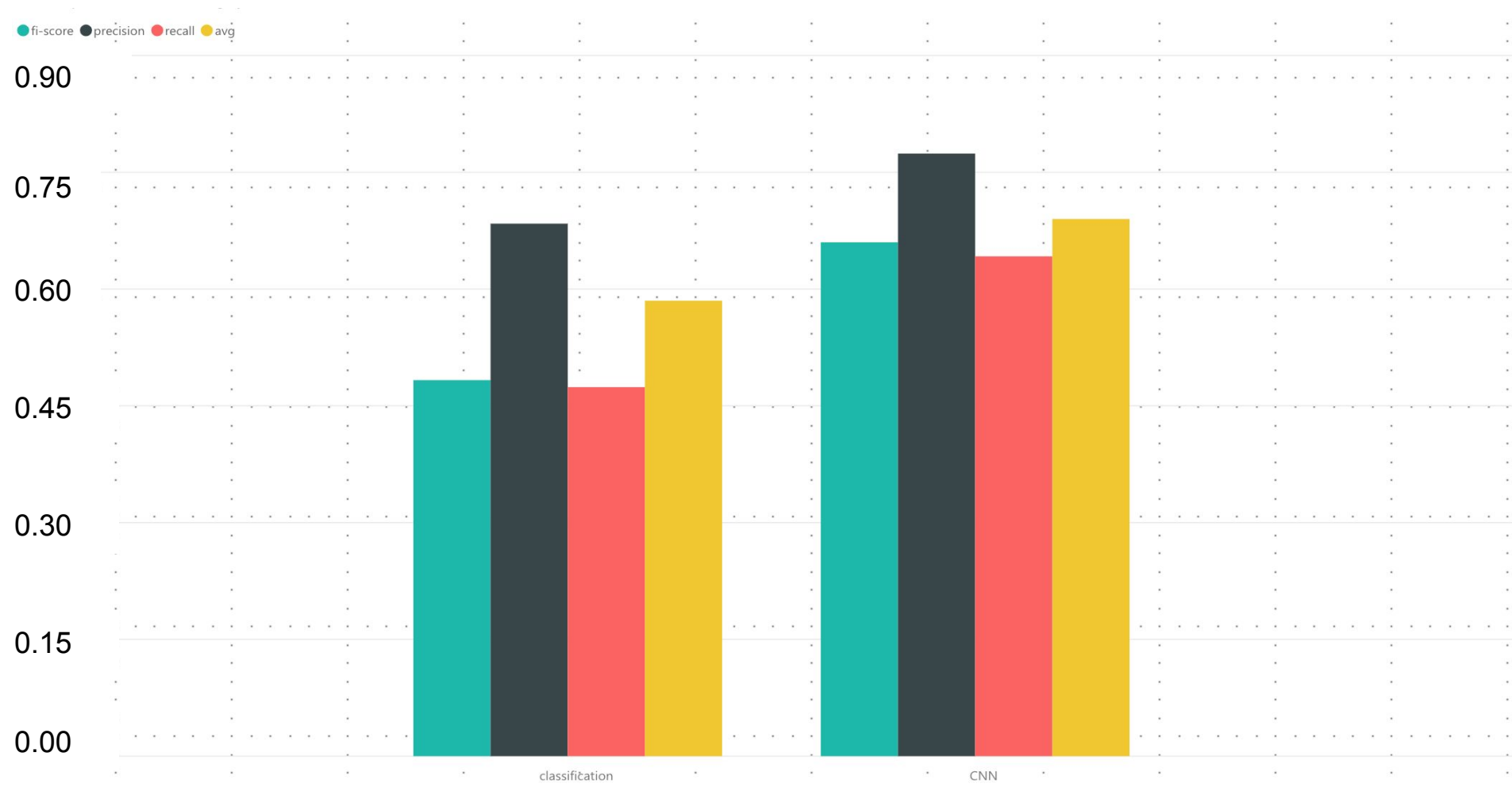
- ❖ Convolutional Neural Works

The performance of classifying label:

	precision	recall	f1-score	support
1	0.85	1.00	0.92	423
2	0.90	0.97	0.88	77
3	0.83	0.23	0.34	0
avg / total	0.86	0.73	0.71	500

1 -> Data Scientist 2-> Sr. Software Engineer 3-> Vice President

Precision, f-score, recall



Convolution Neural Network

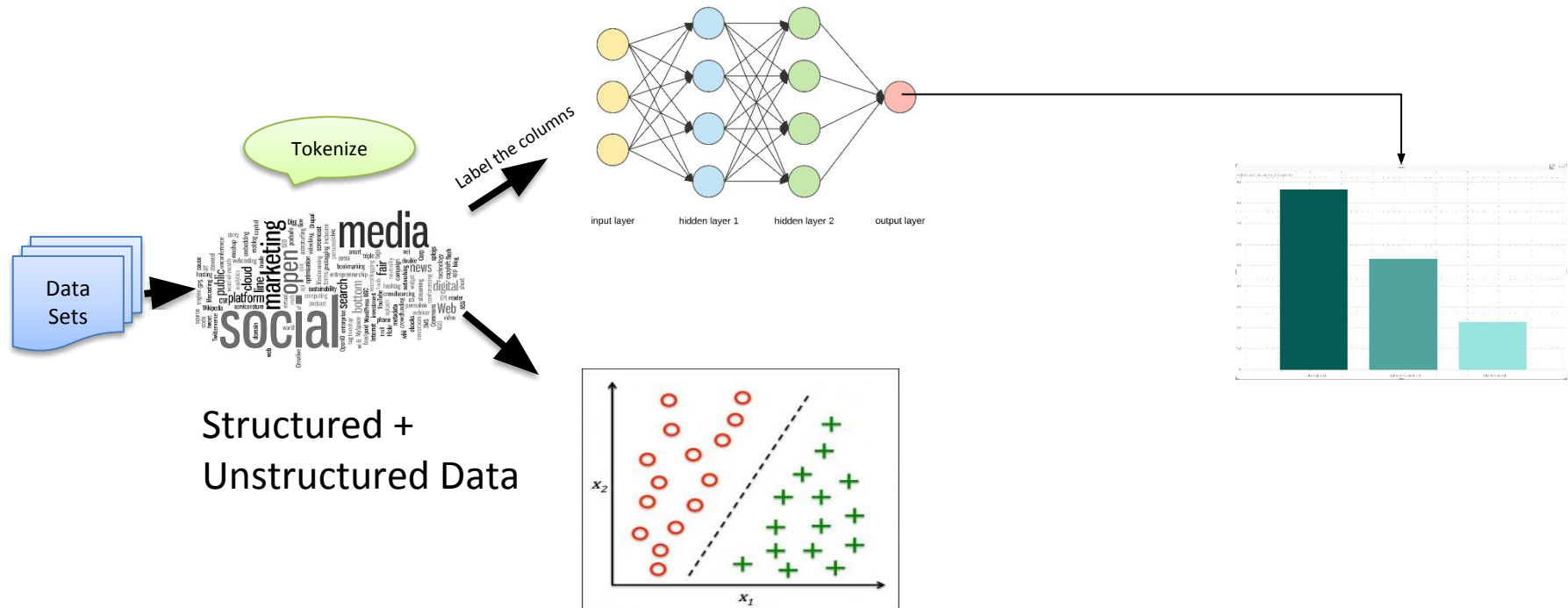
(3, 28026)

'I'm a Vice President / Fashion designer of my own company. I design clothing and fashion ranges.' => 3

'Developed client module to be used in Windows environment in C++, MFC, COM, TCP/IP Sockets' => 2

'Designed and developed a demonstration system that includes all of the options and security features that are available' => 1

Convolution Neural Network

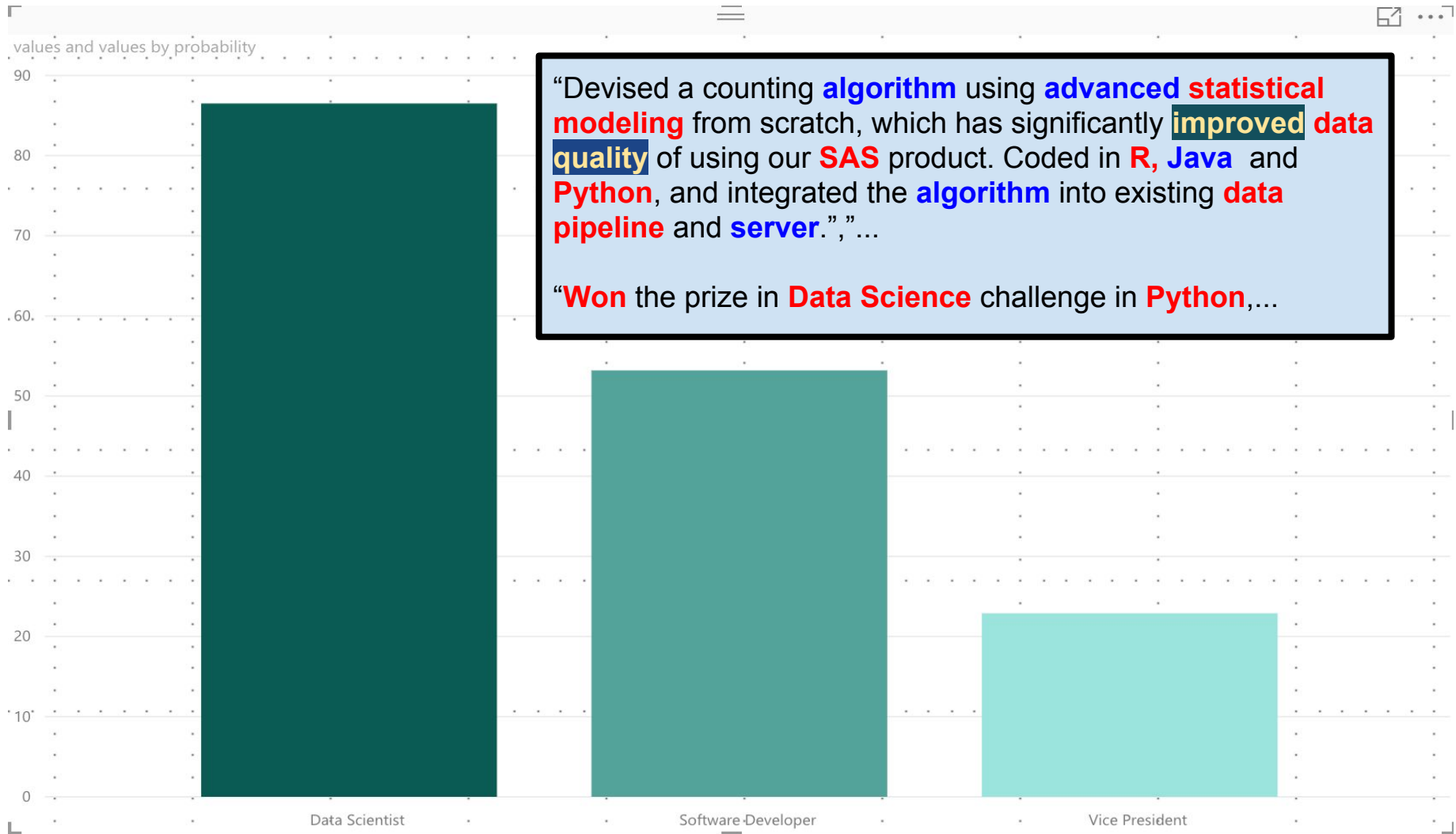


Naïve Bayes Classification Model

Chances of input resume to be a Data Scientist (CNN on job description) are:
84.10631807763958
Chances of input resume to be a Project Lead (CNN on job description) are:
52.37256176994119
Chances of input resume to be a Vice President (CNN on job description) are:
23.521120152419222



Results

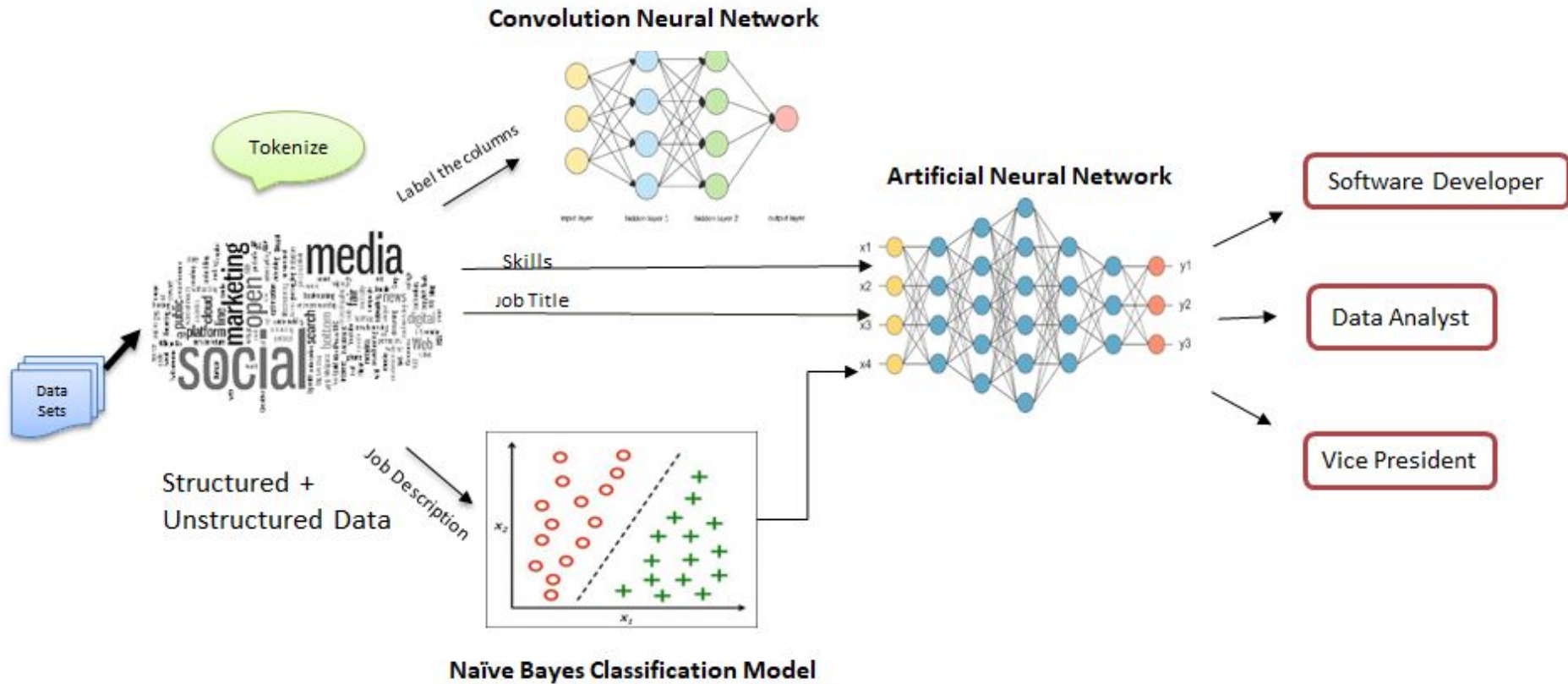




Conclusion & Future Work(1/3)

- ❖ Collecting more data from different profiles
- ❖ Adding other attributes like education, previous titles and many other in prediction.
- ❖ When the data samples get increased, the CNN training time increases as well. So, we will put it into some kind of distributed system.
- ❖ Building career path from current and desired position.
- ❖ Publish a python package (or a API) for third party use.

Conclusion & Future Work(2/3)



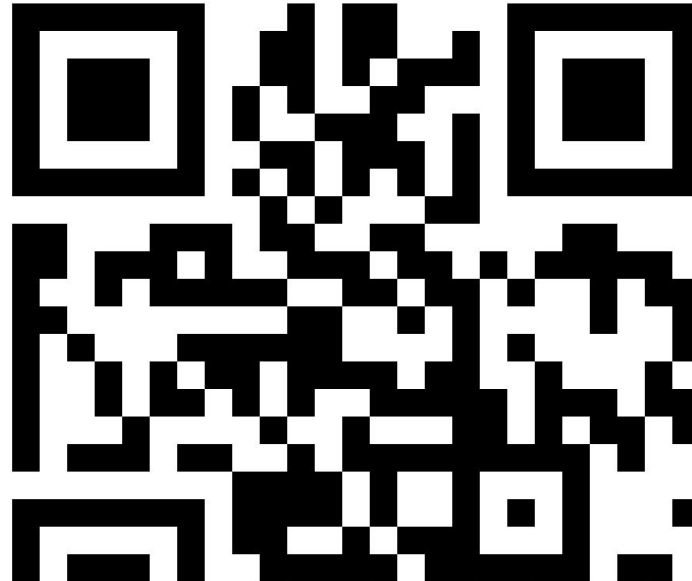


Conclusion & Future Work(3/3)

I am a Software developer, app and I want to be a Construction and building

Chief executive	Social or community	Loan counselor	Civil engineer	Psychologist	Secondary school	Chiropractor	Health practitioner	Police and	First-line	Advertising sales	Credit authorization	Data entry	Pile-driver	Mining machine	Maintenance and	Computer control	Shoe and leather	Jeweler or	Subway, streetcar
General or	Emergency management	Tax examiner	Computer hardware	Sociologist	Special education	Dentist	Licensed practical	Transit and	First-line	Insurance sales	Customer service	Word processor	Operating engineers	Roof bolter.	Maintenance worker.	Extruding and	Shoe machine	Medical, dental.	Sailor or marine
Legislator	Manager, misc.	Financial specialist	Electrical and	Urban and regional	Other teachers	Dietitian	Medical records	Animal control	Animal trainer	Securities	Eligibility interview	Desktop publisher	Drywall installer	Roustabout oil and	Millwright	Forging machine	Tailor, dressmaker	Packaging and	Ship and boat
Advertising or	Purchasing agent.	Computer scientist	Industrial engineers	Agricultural	Librarian	Pharmacist	Misc. health	Security guard and	Gaming services	Sales representative	Hotel, motel, or	Mail clerk or	Glazier	Other extraction	Telecommunications	Cutting, punching.	Textile cutting	Photograph process	Bridge and lock
Marketing and sales	Wholesale or retail	Computer and	Marine engineer	Biological technician	Library technician	Physician or	Other healthcare	Crossing guard	Motion picture	Sales representative	Interview except	Office clerk.	Insulation worker	First-line	Precision instrument	Drilling and	Textile knitting	Semiconductor processor	Parking lot
Public relations	Purchasing agent.	Computer systems	Materials engineer	Chemical technician	Teacher assistant	Physician assistant	Nursing, psychiatric	Transport, security	Ushers, lobby	Model, demonstration	Library assistant	Office machine	Painter, construction	Computer, automated	Wind turbine	Grinding, lapping.	Textile winding.	Cementing and	Service station
Administrative services	Claims adjuster.	Information security	Mechanical engineer	Geological	Other education	Podiatrist	Occupational therapist	Lifeguard	Miscellaneous entertainment	Real estate	Loan interview	Proofreader or copy	Paperhanger	Radio and telecommunications	Coin, vending.	Lathe and turning	Extruding and	Cleaning, washing.	Transport, inspector
Computer and	Compliance officer	Computer programmer	Mining or geological	Nuclear technician	Artist	Audiologist	Physical therapist	Chef or head cook	Funeral service	Sales engineer	New accounts	Statistical assistant	Pipelayer plumber.	Avionics technician	Commercial diver	Milling and	Fabric and	Cooling and	Transport, attendant
Financial manager	Cost estimator	Software developer, applications and systems software	Nuclear reactor	Social science	Designer	Occupational therapist	Massage therapist	First-line	Mortician undertaker	Telemarketer	Order clerk	Electrician or stucco	Plasterer or stucco	Electric motor.	Locksmith and safe	Machinist	Upholsterer	Etcher and	Other transport
Compensation and	Human resource	Compensation and	Human resource	Miscellaneous	Actor	Physical therapist	Dental assistant	Cook	Barber	Door-to-door	Human resources	First-line	Reinforcing iron and	Electrical and	Manufacturing building	Metal furnace	Textile, apparel.	Molder, shaper.	Conveyor operator
Training and	Compensation or job	Subcontract	Subcontract	Misc.	Producer or	Radiation therapist	Medical assistant	Food preparation	Hairdresser, hairstylist	Sales and related	Reception	Agricultural inspector	Roofer	Electrical and	Rigger	Model maker or	Cabinetmaker and bench	Paper goods	Crane and tower
Industrial production	Training and	Administrative	Drafter	Social worker	Athlete, coach.	Recreation therapist	Medical transcription	Bartender	Misc. personal	First-line	Reservation	Animal breeder	Sheet metal	Electrical equipment	Signal and track	Molders and	Furniture finisher	Tire builder	Dredge, excavator
Purchasing manager	Logistics	Engineering	Engineering	Probation officer	Dancer or choreographer	Respiratory therapist	Pharmacy aide	Food preparation	Baggage porter.	Switchboard operator.	Information and	Grader or sorter.	Solar photovoltaic	Electrical home	Helper - installer	Multiple machine	Model maker and	Helper-production	Hoist and winch
Transport, storage.	Management analyst	Computer network	Computer surveying or	Social or human	Musician or singer	Speech-language	Veterinary assistant	Counter attendant	Tour and travel	Telephone operator	Cargo and freight	Miscellaneous agricultural	Structural iron and	Security and fire	Other installer	Tool and die maker	Sawing machine	Production worker.	Industrial truck and
Farmers, ranchers.	Event planner	Computer occupation	Agricultural or food	Misc. community	Entertainment	Exercise physiology	Phlebotomist	Waiter or waitress	Transport, attendant	Communication equipment	Courier	Fishers and	Helper, construction	Aircraft mechanics	First-line	Welding, soldering	Woodworking	Supervisor of	Cleaner
Construction manager	Fundraiser	Actuary	Biological scientist	Clergy member	Announcer	Therapist all other	Misc. healthcare	Food server.	Child care	Bill and account	Dispatcher	Hunter or trapper	Automotive structure	Aircraft structure	Heat treating	Woodworking	Aircraft pilot or	Laborer and	Machine feeder or
Education administrator	Market research	Mathematician	Conservation scientist	Director, religious	News analyst.	Veterinarian	First-line	Dining room.	Personal and home	Billing and	Meter reader.	Forest and	Plasterer or	Automotive glass	Electrical electronics	Lay-out worker.	Power plant	Air traffic	Shuttle car
Engineering manager	Business operation	Operation: research	Medical scientist	Religious worker.	Public relations	Registered nurse	First-line	Dishwasher	Recreation and	Bookkeeping accounting	Postal service	Logging worker	First-line	Engine and other	Structural metal	Plating and	Stationary engineer	Flight attendant	Packer or packager.
Food service	Accountant or	Statistician	Life scientist	Lawyer, judges.	Editor	Nurse anesthetist	First-line	Host or hostess.	Residential advisor	Gaming cage	Postal service	First-line	Boilermaker	Brickmason blockmason	Septic tank	Misc. vehicle	Butcher and other	Crushing, grinding.	Tank car, truck, or
Funeral service	Appraiser of real	Miscellaneous	Astronomer	Judicial law clerk	Technical writer	Nurse midwife	Supervisory protective	Fire fighter	First-line	Payroll and	Postal service	First-line	Brickmason blockmason	Septic tank	Misc. vehicle	Butcher and other	Crushing, grinding.	Tank car, truck, or	Material moving
Gaming manager	Budget analyst	Architect except	Atmospheric or space	Judge, magistrate	Writers and	Nurse practitioner	Fire fighter	First-line	First-line	Procurement clerk	Shipping, receiving	Stock clerk or	Carpet, floor.	Misc. construction	Food and tobacco	Print binding	Cutting worker	Motor vehicle	Material moving
Lodging manager	Credit analyst	Surveyor or	Chemist or legal	Miscellaneous legal	Broadcast and sound	Clinical laboratory	Bailiff, correction	Janitor or	Cashier	Financial clerk.	Stock clerk or	Carpet, floor.	Misc. construction	Food and tobacco	Print binding	Cutting worker	Motor vehicle	Material moving	
Medical and	Financial analyst	Aerospace engineer	Environmental scientist	Miscellaneous legal	Broadcast and sound	Clinical laboratory	Bailiff, correction	Janitor or	Cashier	Financial clerk.	Stock clerk or	Carpet, floor.	Misc. construction	Food and tobacco	Print binding	Cutting worker	Motor vehicle	Material moving	

QR code for Demonstration.



Welcome to Resume Analysis. Project of BIA-660-B Team 16.

Enter the Job Descriptions from your resume:

- Constructed batch job scripts in ETL (transformations and data validation phase) for Informatica tool to ensure accuracy and efficiency in analysis reports. (For investment banking clients)
- Developed, deployed and tested OBIEE-12c-cloud reports and dashboards according to clients' requirements to improve and enhance product's sale in next quarter. (For automobile clients)
- Led a team for a customized, cross-platform e-commerce web-app using Pthon, R, SAS, Hadoop.

Submit

Your Chances of becoming a Data Scientist are: 74.49060211295702 %



STEVENS
INSTITUTE *of* TECHNOLOGY
THE INNOVATION UNIVERSITY®

stevens.edu

programous.com

mehulmistry.com