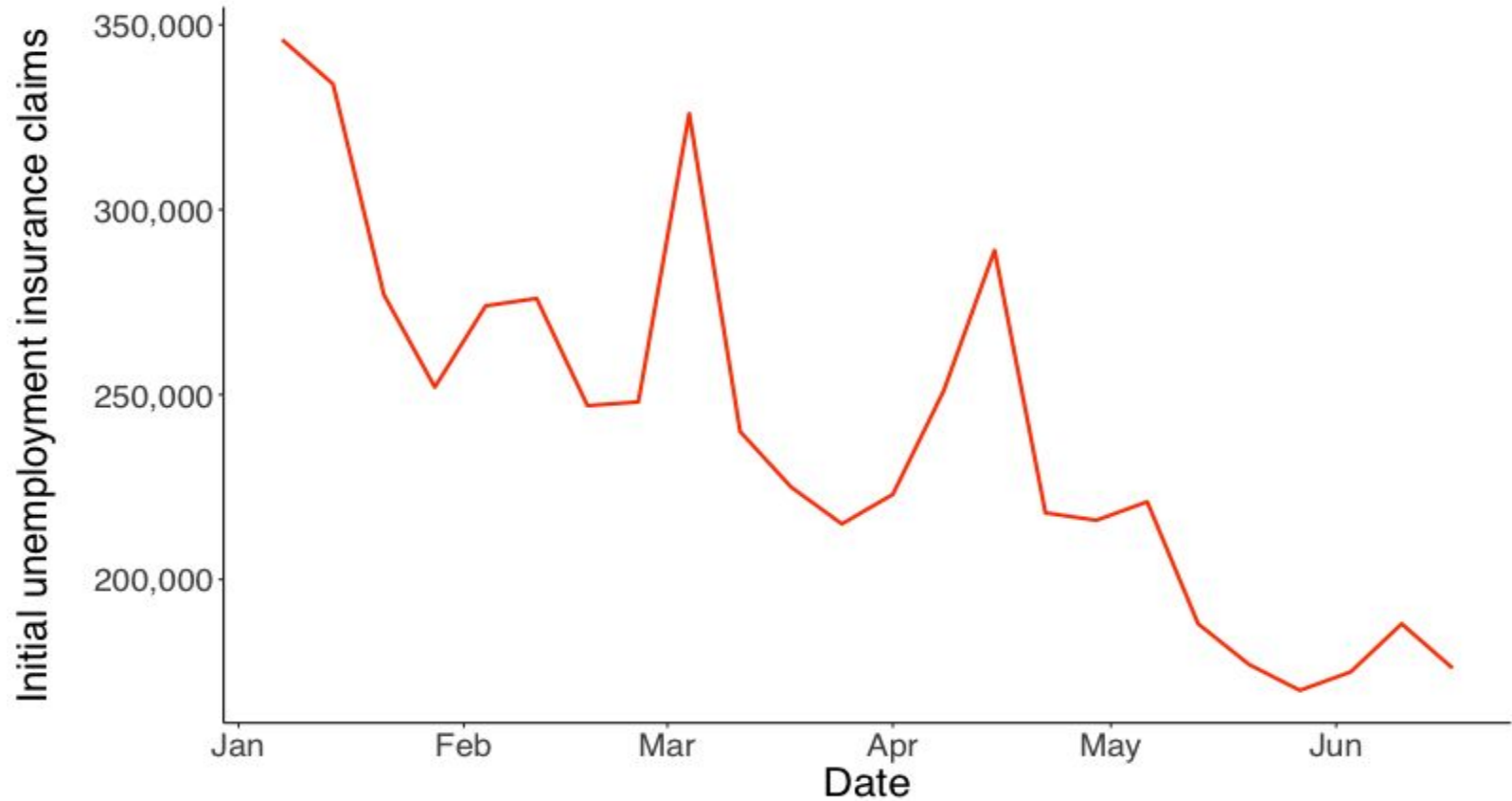




Tech Job Recommender

Nolan Arendt, Zachary Brown, Albert Frantz

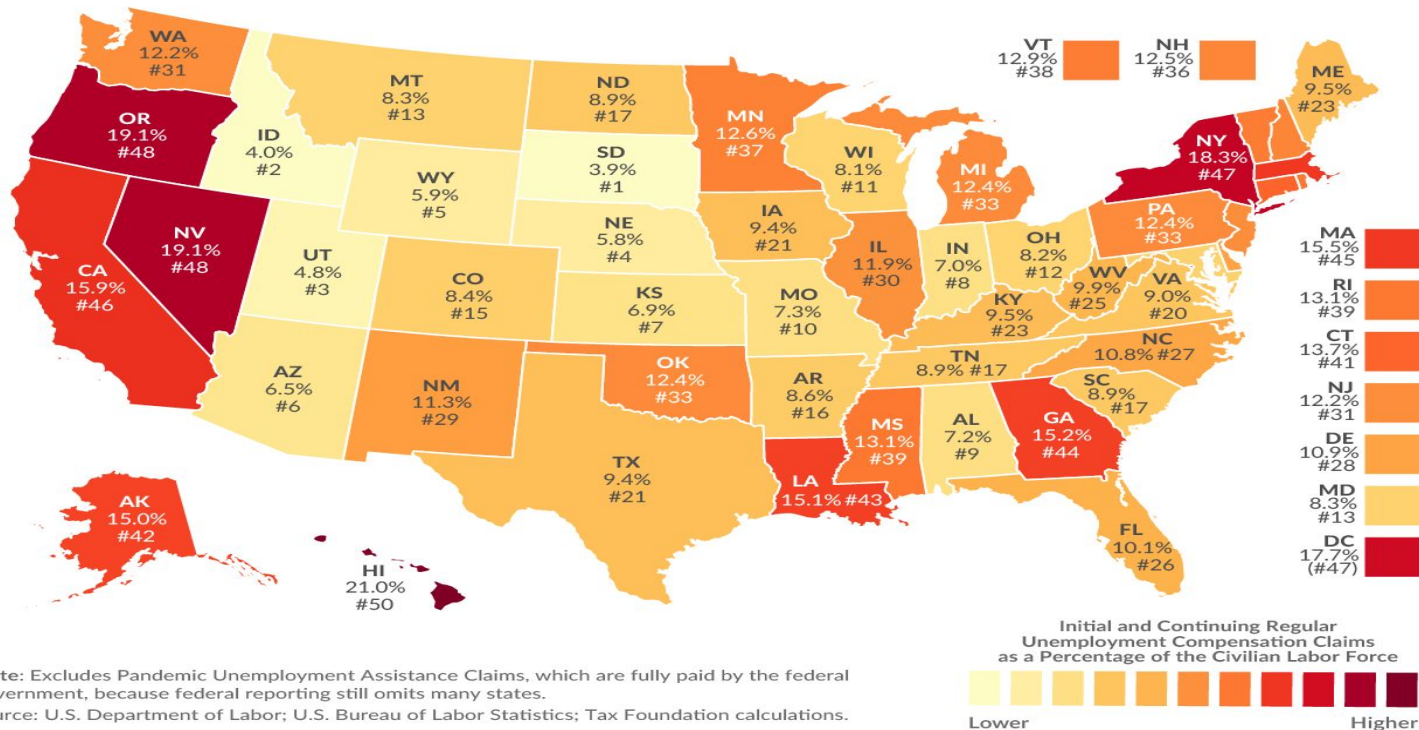
Unemployment Anomaly of 2020



Unemployment Anomaly of 2020

Unemployment Compensation Claims as a Percentage of the Civilian Workforce

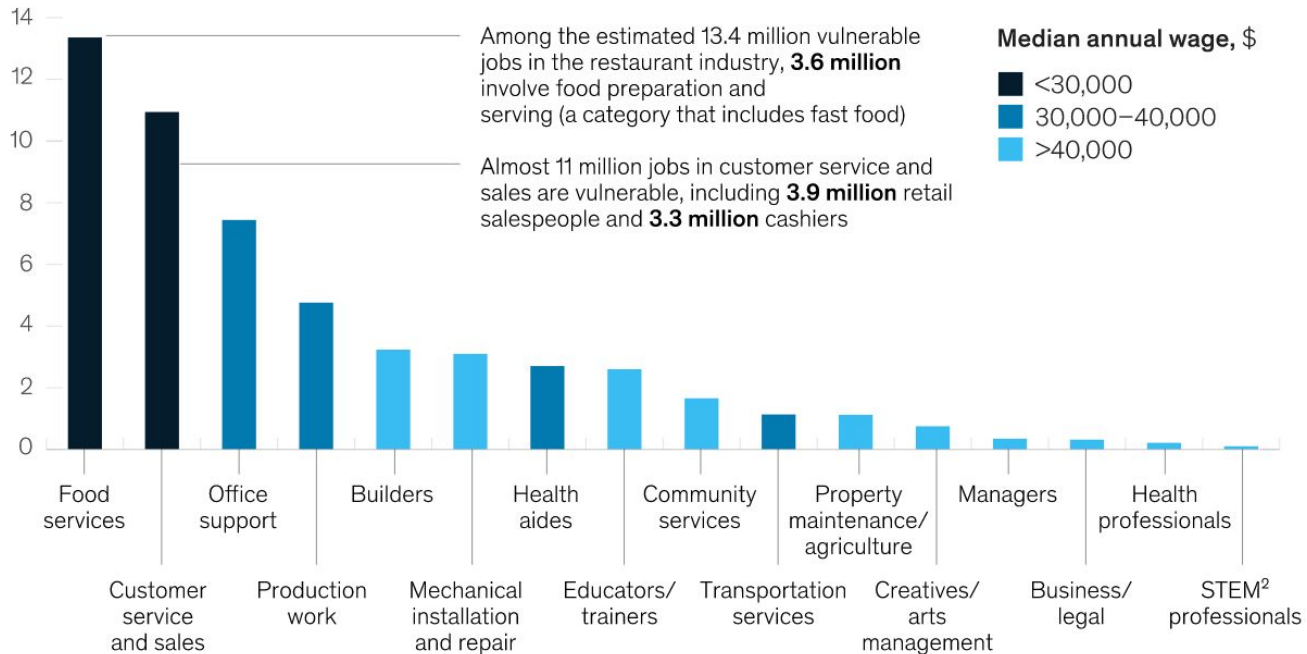
Initial and Continuing Regular Unemployment Claims as a Percentage of the Civilian Labor Force, June 20, 2020



Unemployment Anomaly of 2020

Forty-six percent of vulnerable jobs are in food service, customer service, and sales.

Vulnerable jobs,¹ by occupation, millions





Dead End

↖ NEXT EXIT

**Career
Change**

NEXT EXIT ↗

“Changing careers is like driving down the highway at 200 miles per hour and wanting to get off at the next exit. You need to slow down, signal, and move yourself towards the ramp with plenty of time or you’ll miss it ... In short, changing careers ... takes planning and patience.”

By J.T. O'Donnell

MISSION STATEMENT

Make the transition into the tech industry less intimidating by matching unemployed individuals to their best fit in tech.



MISSION STATEMENT CONTINUED

Short Term:

Help those unemployed due to COVID-19 find a matching tech title to apply for.

Long Term:

Help guide people undergoing a career change to the tech industry to optimize their job search.

Core Values

- Assist those affected by COVID-19 unemployment to find their perfect fit in the tech sector.
- Help guide people undergoing a career change to the tech industry to optimize their job search.
- Optimize job application efficiency.

01

02

03

04

05

PROJECT OVERVIEW

DATA CLEANING

- Jobs data set
- Linkedin data

RECOMMENDER SYSTEMS

- SpaCy
- Cvec/Tfidf
- KNN

BEST RECOMMENDER

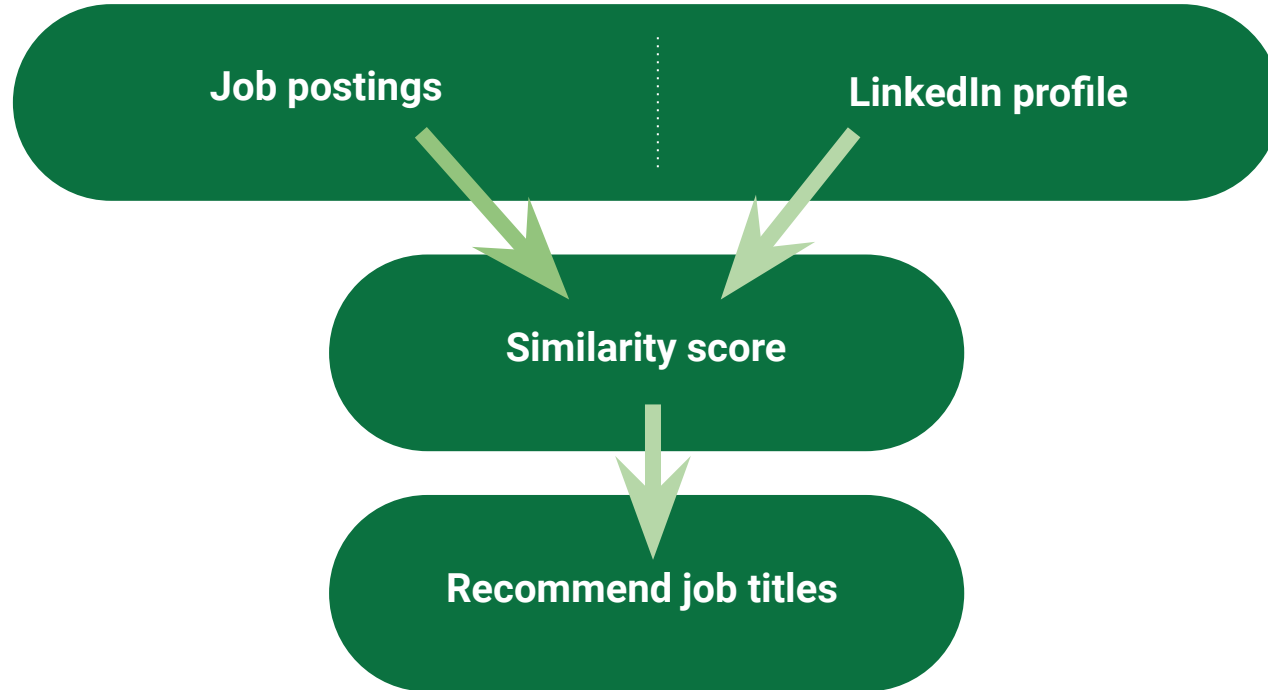
FUTURE ANALYSIS

TABLE OF CONTENTS

PROBLEM STATEMENT

Determine the best approach to building a recommender system that pairs job seekers with the job title they match best, based on a job seeker's LinkedIn profile.

PROJECT OVERVIEW



DATA CLEANING

01

Removing 'fluff'

Special characters
Duplicate phrases
Phone numbers and emails
Duplicate posts

02

Categorizing job titles

categorized job titles into 13 groups ie: developer, engineer, Data position etc.

03

Categorizing location

categorized location into four categories:
Northeast, Midwest, Southern, Western

04

Categorizing job type

categorized job type into three employment types:
Full Time, Contract, Market Dependent

05

General Cleaning

THE FINAL DATA

	job_title	job_description	skills	location
0	Analyst	If you are an EDI Analyst with experience, ple...	edi trustedlink as2 van	Northeast United States
1	Developer	Web/Backend Dev (Ruby on Rails)Seattle, WAEith...	ruby on rails mobile web aws javascript html5	Western United States
2	Manager	This nationally recognized Microsoft Gold Part...	microsoft dynamics ax project manager - toront...	Western United States


JOB TITLES

Analyst - 1900	Director-146	Manager - 1390
Administrator - 899	Developer - 5274	Technician - 343
Consulting - 592	Engineer - 4118	Architect - 831
Data Position - 283	Programmer - 321	Support - 343
Designer - 221		


- More overlap than we expected between job titles (Analyst/Data position, Developer/Programmer).
- We stuck strictly to the wording used in the job postings.

LINKEDIN DATA


Experience

**Data Science Fellow**
General Assembly · Full-time
Oct 2020 – Present · 3 mos
Denver, Colorado, United States

- Participated in a full-time 3-month data science course.
- Completed various in-class projects, hackathons, programming challenges, and personal projects focused on applying data science concepts to real-world problems while learning industry best practices.

**Assistant Teacher**
Horizons National · Seasonal
Jul 2017 – Aug 2019 · 2 yrs 2 mos
Greater Denver Area

- Assisted the lead teacher in the continuation of summer education for academically struggling low-income students.
- Directed volunteers, interns, and students of kindergarten, first grade, and ninth grade classes.
- Multi-tasked by leading class, watching for the safety of students and preparing daily n...see more

**Classroom Intern**
Horizons National · Internship
Jul 2015 – Aug 2016 · 1 yr 2 mos
Greater Denver Area

- Assisted the lead and assistant teachers in the daily management of a fifth-grade classroom.
- Facilitated and organized the graduation of the fifth-grade class.

Skills & endorsements

Take skill quiz

R · 1
Chase Wolf-Tinsman has given an endorsement for this skill

Python · 8
Endorsed by 6 of Albert's colleagues at General Assembly
Passed: LinkedIn Assessments

Data Analysis · 7
Endorsed by 5 of Albert's colleagues at General Assembly

Show more

About

I am a detail-oriented data scientist that uses data as a language to understand the problems we face, from making cities more efficient to creating art. A question raiser who loves to apply data in new and unconventional ways. With non-profit experience, I am driven by the constantly evolving needs of those around me to use data to improve...

	Titles	Skills	Summary	Projects	Certifications	Education
Albert Frantz	Data Science Fellow, Assistant Teacher, Classroom Intern	R, Python, Data Analysis, Econometrics ...	I am a detail-oriented data scientist that uses data as a ...	Identifying the Relationship Between ...	Tableau 2020 A-Z: Hands-On ...	Bachelor of Arts - BA

COUNTVECTORIZER MODEL

Collection of words from user profile data and job descriptions.

Counts how many times a word is used in both profile data and job descriptions.

Stores the number of occurrences for each word.

Comparison of words from profile data and job descriptions to calculate a cosine similarity score.

<u>Name</u>	<u>Recommended Position</u>	<u>Recommender Score %</u>
<i>Albert</i>	<i>Data Position</i>	<i>19.1%</i>
<i>Nolan</i>	<i>Data Position</i>	<i>20.5%</i>
<i>Zach</i>	<i>Data Position</i>	<i>17.8%</i>
<i>Ye</i>	<i>Director</i>	<i>3.6%</i>

KNN MODEL

Goes through the same steps as countvectorizer but gathered the 75 most closely related job postings

Calculates the frequency of each job title in those 75 most related jobs

<u>Name</u>	<u>Recommended Position</u>	<u>Job Count</u>	<u>Job Match %</u>
<i>Zach</i>	<i>Data Position</i>	<i>37</i>	<i>49.3%</i>
<i>Nolan</i>	<i>Data Position</i>	<i>26</i>	<i>34.7%</i>
<i>Albert</i>	<i>Data Position</i>	<i>26</i>	<i>34.7%</i>
<i>Ye</i>	<i>Developer</i>	<i>20</i>	<i>26.7%</i>

SpaCy

Context

Synonyms and word meaning

Grammer

<u>Zach's Scores</u>
91.589%
91.555%
91.212%

<u>Zach</u>	<u>Nolan</u>	<u>Albert</u>	<u>Ye</u>
<i>Analyst</i>	<i>Developer</i>	<i>Director</i>	<i>Director</i>
<i>Data Position</i>	<i>Programm er</i>	<i>Manager</i>	<i>Manager</i>
<i>Engineer</i>	<i>Architect</i>	<i>Analyst</i>	<i>Analyst</i>

Our Choice: CVEC

- Largest spread in similarity scores
- Fastest to load
- Results consistent with other models

Comparison - Albert's Profile

<u>Model</u>	<u>Job Recommended</u>	<u>Spread</u>
<i>CVEC</i>	<i>Data Position</i>	<i>11.61%</i>
<i>Tf-Idf</i>	<i>Data Position</i>	<i>4.14%</i>
<i>SpaCy</i>	<i>Director</i>	<i>3.66%</i>
<i>KNN</i>	<i>Data Position (34.7%)</i>	<i>-----</i>

FUTURE ANALYSIS

Implement into Web Application

- Available for any user wanting to test their LinkedIn profile information

Applications to Other Industries

- LinkedIn, Glassdoor, USAJobs, etc

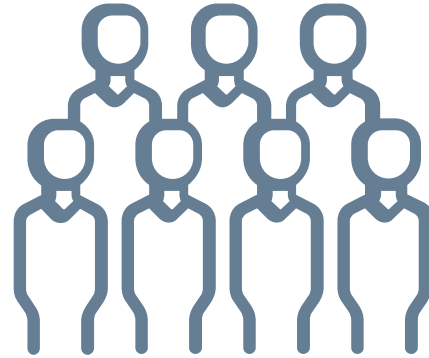
Personalized Job Recommendation List

- Sector specific

Compatibility with Resume

Streamline Cleaning Process

- Pipeline to process new jobs data



THANKS

Does anyone have any questions?
Follow the project updates



- Reddit Visualization Unemployment Claims in United States
<https://www.reddit.com/user/DorsaAmir/>
- Happy professional picture:
https://www.thetaylormethod.com/wp-content/uploads/2016/12/New_Mindset_black_professional_happy-min.jpg
- Happy professional picture 2:
https://cdn.hrpayrollsystems.net/wp-content/uploads/2015/03/Dollarphotoclub_77841384.jpg
- Covid Map of States:
<https://taxfoundation.org/unemployment-insurance-claims/>
 - Unemployment by sector
<https://www.mckinsey.com/industries/public-and-social-sector/our-insights/lives-and-livelihoods-assessing-the-near-term-impact-of-covid-19-on-us-workers>
- Presentation template by **Slidesgo**
- Icons by **Flaticon**
- Infographics by **Freepik**
- Images created by **Freepik**

CREDITS