2020 HIMCM PROBLEM A

HIMCM GROUP 10656

1. Summary

The uncertainty caused by the COVID-19 pandemic has led many to begin making organized plans for the future. Among high school students, this involves looking to the summer of 2021 for crucial work experience amidst the pandemic. However, the job market is often intimidating through its sheer volume and variety, especially for high schoolers— many of whom have no experience with job-searching. In this paper, we develop a comprehensive model matching high schoolers with suitable jobs based on several inputs set by the job seeker. We extend this model to unpaid internships, another popular summer work opportunity for high school job seekers.

We first define several factors which delineate the most important priorities of a high school job candidate. These are, namely, COVID-19 risk, maximum physical exertion, wage, working hours, and fields of interest. The user's input data is matched to a database of work opportunities, each with their own sets of data to illustrate the work experience. Combining user and job data in an ultimate mathematical formula yields a calculated score of suitability corresponding to every user-job pairing, with the ten most suitable jobs by score returned for the user. A similar formula is developed for unpaid internships, with a heavier emphasis on user interests.

Both models are implemented in Python, allowing for easy testing of large and diverse data. The models are tested on 17 fictional high school students with varying backgrounds: 10 searching for jobs and 7 searching for internships. Varying backgrounds allow for both extreme and moderate cases to be tested and evaluated. Holistically, the test data would be completely representative of all high school job seekers, given the criteria.

We present our model as a simple and navigable online website. Job seeker input is gathered through a short questionnaire, establishing the user's priorities, and our mathematical model is then run to return the best work opportunities for the user, based on their answers to the questionnaire. This presentation allows summer jobs and internships to be matched to high school opportunity seekers with ease, simplifying the chaotic job market into an organized, individualized list, and establishing a plan for the future during the uncertain present.

Contents

1. Summary 1
2. Introduction 2
2.1. Background 2
2.2. Problem Interpretation 2
2.3. Assumptions 3
3. Developing the Model 4
3.1. Variables 4
3.2. Developing the Data 6
3.3. A Mathematical Formula for Jobs 7
3.4. A Mathematical Formula for Internships 10
4. Testing the Model 10
4.1. Sensitivity Analysis 13
5. Presenting the Model 13
6. Strengths and Weaknesses 15
7. Conclusion 16
References 17
8. Appendix I: Python Code for the Job Model 17
9. Appendix II: Python Code for the Internship Model 19
10. Appendix III: Full Job Recommendations for All Test Cases 20
11. Appendix IV: Full Internship Recommendations for All Test Cases 21
12. Appendix V: Full job and test database 22
13. Appendix VI: Full internship and test database 23

2. Introduction

2.1. **Background.**

As the end of 2020 approaches, high school students must plan ahead and search for work opportunities for the upcoming summer. This is easier said than done. Each student possesses a unique and diverse pool of interests, preferences, and needs which are suited to a myriad of different summer jobs. Furthermore, most high school students— overwhelmed by the seemingly infinite number of choices while unsure of where to find specific job opportunities— are unaware of which jobs are best for them, prompting a need for an automated model to help them make a decision.

The main problem lies in the countless number of factors which can be considered for each potential job. This only increases with the current COVID-19 pandemic. To avoid confusion, these factors must be dealt with by an automated algorithm or model, in which the most significant interests and priorities must be extracted and processed to determine the "best" job.

2.2. **Problem Interpretation.**

Our objective is to create an algorithm which determines the best job match for a high school student in the summer of 2021 based on several aspects of the job seeker and the jobs themselves. We interpreted the problem as three fundamental questions:

(1) What jobs for high schoolers will be available in the summer of 2021, and what will the job experiences entail?
(2) What aspects of a summer job will an individual deem important? How can they be determined or quantified? What aspects of an individual will an employer deem important?

(3) How can we evaluate the many opportunities for high school summer work opportunities and find those which best fit an individual based on their interests, preferences, and priorities?

The solutions to the first and second questions define the factors needed to determine a suitable job match. Answering the third question combines these factors into a meaningful model which can be quickly used to determine the best fit. Keeping in mind the target audience, the model should be simple and presentable. We can also extend the model to search for unpaid internships, which involves a similar process.

2.3. **Assumptions.**

It is necessary to make a few preliminary assumptions, so that a clear and logical solution to be outlined:

(1) Assumption 1: Demographics

The model will only apply to high schoolers 14 years of age and above who are American citizens residing within the United States.

Justification: The bulk of occupational data and statistics available to us were collected from the United States of America. The vast majority of American citizens attending American high schools reside in the United States and would therefore be considering job opportunities within the United States. American federal child labor laws prohibit non-agricultural work for minors below the age of 14. Our model will only consider and encourage legal employment opportunities, as stated in Assumption 3. Thus, 14 years is used as the threshold age for a significantly wide range of legal employment opportunities to open up.

(2) Assumption 2: Definition of a Summer Job

A summer job is defined as a position where one is legally employed and paid for repeated periods of labor over the course of a summer, while high school is not in session. This includes standard paid jobs and paid internships, but not unpaid internships and volunteering.

Justification: This definition of a job is widely accepted internationally. American high schools are not in session during most of the summer.

(3) Assumption 3: Extent of Legality

All jobs are compliant with American federal labor laws, but not necessarily state or local laws.

Justification: Labor laws vary to an extremely high degree amongst states, and to an even higher degree locally. To offer a more uniform picture for all American high schoolers, we set the universal standard at the federal level. This means that unreported "under the table" jobs are not considered in this model, as well as any violations of age requirements.

(4) Assumption 4: Qualifications

High schoolers cannot hold any degree of education at or above a high school diploma, and all employers follow only the U.S. Bureau of Labor Statistics Occupational Outlook Handbook (OOH) in accepting applicants' qualifications.

Justification: Fewer than 3% of high schoolers graduate early [8], so it is safe to assume that a high schooler using the model will not hold any formal degree of education. Although 18 year-olds can have high school diplomas, which open up the job market, we found that most of these jobs, such as being a police officer or commercial pilot, require additional training after graduation and could therefore not be feasibly held the summer after graduation. For simplicity and accuracy, we assume that no high schooler using our model holds any higher educational degree. We base our data off of OOH, as it is a definitive federal governmental database of labor statistics. This has two important ramifications regarding qualifications. First, all job categories which require a high school diploma, according to the OOH, are restricted from the model. Second, the remaining job categories will not consider any qualifications in accepting applicants.

(5) Assumption 5: Job Categories
All jobs within our given categories involve the same work, physical intensity, hours, and wages.

Justification: Specific aspects of jobs will vary widely based on employer and location within the United States. These aspects cannot be specifically identified even at the state or local level. For simplicity, we assumed that nationwide metrics would dictate the relative positioning of job categories in terms of our chosen factors.

(6) Assumption 6: COVID-19 Safety Guidelines
All in-person jobs follow all COVID-19 safety guidelines set by their respective states in November 2020.

Justification: We unfortunately project that the COVID-19 pandemic will still be widespread in the summer of 2021. Currently, many jobs take place in-person during the pandemic with added safety restrictions. We assume all jobs are legally compliant with all of these restrictions to establish a standard. This assumption allows the COVID-19 safety of a job to be evaluated solely based on the amount of contact on the job, which could vary widely based on location and regulation, but would not vary significantly relative to other jobs.

(7) Assumption 7: Consistent Work Schedule
For all jobs, every day of work will consist of the same relative number of hours per day.

Justification: Typically, working hours can vary on a daily and weekly balance, especially for jobs with greater demand at certain times and days (e.g. restaurant cook). For simplicity, and to create a standardized time frame for working hours, we assume the average working hours per work day.

3. Developing the Model

3.1. **Variables.**

We list the variables at play in our model below:

(1) R_c is the job's COVID-19 risk rating, on a scale from 1 to 4, with 1 being the lowest risk and 4 being the highest. Guidelines from OSHA [6] were used to score each job. The user input R_i, measured on the same scale, reflects the necessity of the summer job, based on how much of a COVID-19 risk the user is willing to accommodate for the sake of working.

(2) Wages (measured in United States Dollars per hour) are the primary indicator of how much money is earned from a job. In our model, we draw entry-level wages from a job database,

which reflects nationwide data from 2018-2020 [7]. We classify these into three wage brackets, based on the hourly wage of the job:

(1) \$8-9, (2) \$9-11, (3) \$11+. The job wage value W_c is then $1, 2$, or 3, depending on which wage bracket the job falls under. The user input W_i reflects the level of priority the user places on earning money through the summer job, inputted as the desired wage bracket the user would like to work for.

(3) Hours (measured in working hours per day) are classified into three brackets:

(1) 1-2 hour jobs, (2) half day jobs, or (3) full day jobs. The jobs each have two corresponding hour variables, H_m and H_M, each being a number form 1 to 3. These represent the minimum and maximum time brackets the job generally provides.

The user input H_i is the desired time bracket they desire to work for. It reflects the user's level of commitment to the job, versus other summer activities and commitments such as recreation or academics.

(4) We rank physical intensity on a four-point scale. A score of 1 means the job is completely sedentary and requires virtually no physical movement (e.g. virtual tutor). A score of 2 means the job requires basic standing or minimal movement (e.g. cashier). A score of 3 means the job requires walking or moderate movement (e.g. waiter), and a score of 4 means the job requires strenuous physical activity (e.g. construction worker). The value P_c shows how much physical activity the job requires.

The user input P_i is the maximum amount of physical activity the user is willing to exert, which reflects the range of comfort the user has with varying levels of physical intensity. This may be affected by physical disabilities or injuries.

(5) Each job has a legal minimum age requirement A_c of either 14, 16, or 18. This variable determines which jobs a student can legally apply to, and which students an employer can legally hire. The user input A_i is simply the age of the user.

(6) Each job was tagged with a number of the following seven interest groups, which were adapted from several of the most popular fields of youth summer employment as in [3]:

(a) Food
(b) Service
(c) Sales
(d) Art
(e) Technology
(f) Manual Labor
(g) Healthcare

A synthesis of the data shown in [3] is shown in Figure 1.

The interest groups submitted by the user are converted into a binary string I_i, where a 1 in the kth bit denotes an interest in the kth field, while a 0 denotes that the user is not interested. This is matched to the associated binary string of each job (I_c), to determine how well the job fit the specific interests of the user.

This prioritizes jobs that applicants would enjoy working at, with a benefit to mental health and job performance.

For internships, the list of interests is different and more specific, since internships tend to be more scientific and specialized. The list of interests is as below, inspired by the most

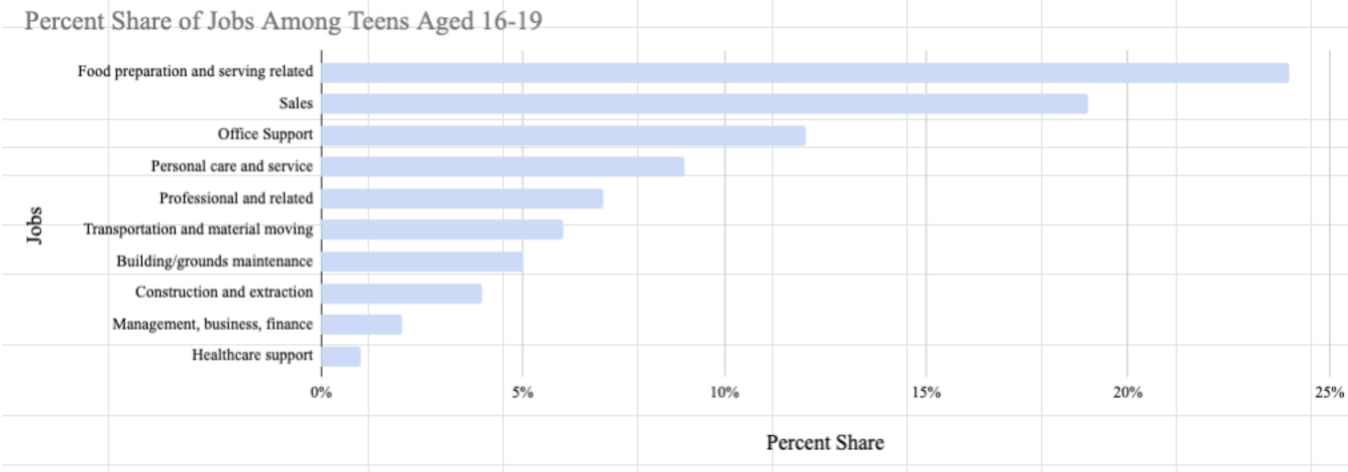


FIGURE 1. The data in [3]. The list of seven interests was drawn from this list of popular jobs.

popular internships listed at [4]:

(a) Education
(b) Life Sciences
(c) Physical Sciences
(d) Mathematics
(e) Data Analysis
(f) Engineering
(g) Finances
(h) Communications
(i) Service
(j) Art
(k) Technology

(7) The variable E_i accounts for the user's preference of gathering experience (interest) or earning pay. The value is set to one if the user values wages or zero if the user values interest.

3.2. **Developing the Data.**

Using Assumption 4 - that applicants were not allowed to hold a high school diploma- we filtered the U.S. Bureau of Labor Statistics Occupational Outlook Handbook (OOH) for job categories which fit this criteria [7]. This yielded a list of 108 occupations with the only criteria being that no formal education was required.

From this list, we expanded certain categories. For example, the category "Ushers, Lobby Attendants, and Ticket Takers", was expanded to three different jobs, keeping in mind that each of these, while similar in type of work, are different in terms of COVID-19 risk and physical intensity. We also removed categories, specifically categories which we found redundant. For example, the category "Miscellaneous entertainers and performers, sports and related workers" too broad a category to expand, and decided that two existing categories, "Musician/Singer" and "Artist" would sufficiently cover the feasible occupations. This resulted in a list of 107 occupations which high schoolers could consider.

After this adjustment, we referenced federal child labor laws in conjunction with the OOH [2]. Each of the jobs was assigned to an age bracket based on how old the applicant had to be to complete typical tasks required by the jobs. Federal law has preexisting age requirements for certain jobs, each group having an age minimum of 14, 16, and 18 years. From this list, we selected a diverse pool of 36 specific jobs representative of each age bracket (each with its legal age) and a diverse array of our 7 interest categories.

For every job, we assessed the COVID-19 transmission risk R_C on a discrete scale of 1 (highest risk) to 4 (lowest risk), which was adopted from a Department of Labor and Department of Health and Human Services "Guidance on Preparing Workplaces for COVID-19" handbook [6]. Seeing as very few jobs available to high schoolers met the "very high" or "high" risk level established by the handbook, we adjusted the scale within the "medium" and "low" risk levels. We assigned all virtual and remote jobs to a R_c of 1, and jobs with the most exposure (such as waiters/waitresses) to a maximum R_c of 4.

We then assessed the hourly wage W_C for each job, drawing from nationwide entry level wage data like OwlGuru as well as median wage data from O*NET (from [5], [1] respectively). We were able to group the jobs into the three general wage brackets outlined in Section 3.1.

We then assessed the physical demand P_c for each job, using our self-developed scale outlined in Section 3.1. Each job's physical demand was derived from the job descriptions listed at [1]. All virtual and remote jobs were graded as less physically demanding relative to its in-person variant.

Finally, we developed a range of work hour brackets between H_m and H_M to represent the time commitment to the job. Similarly, each job's time commitment was derived from the job descriptions listed at [1].

With all of these variables determined, we were left with a thorough profile of 36 jobs, each with justified assessments for age requirement, COVID-19 transmission risk, hourly wage, physical demand, work hours, and set of interests. This profile is shown in Appendix V. A similar process was done to curate a list of 25 internships, as shown in Appendix VI.

3.3. **A Mathematical Formula for Jobs.**

We attempt to create a mathematical formula to assess how fit a certain job is for an individual. This formula works by giving each job j a score S_j. Higher scores mean the job is more fit to the individual. First, we define seven terms, which will then be multiplied together to get the final score.

$$a_1 = \left(\frac{\left(\frac{|R_i - R_c + \frac{1}{2}|}{R_i - R_c + \frac{1}{2}} \right)}{2} + \frac{1}{2} \right) (1.04)^{R_i - R_c}$$

The first term, a_1, is the COVID-19 risk term. As stated in Section 3.1, each job has a COVID-19 risk rating from 1 to 4, depending on the risk of contracting the disease at the workplace. This is R_c. The user then inputs a number R_i that is the maximum COVID-19 risk rating they are willing to endure.

The first term in the product is 1 exactly when $R_i \geq R_c$, and 0 otherwise. In other words, if the job has a higher risk than the user is willing to endure, its total index is 0 and the job is thrown out of consideration. Otherwise, nothing happens. To see this, note that

$$\frac{|R_i - R_c + \frac{1}{2}|}{R_i - R_c + \frac{1}{2}}$$

is -1 if and only if $R_i - R_c \leq -\frac{1}{2}$. Since R_i and R_c are both integers, this is equivalent to $R_i < R_c$. Dividing by 2 and adding $\frac{1}{2}$ yields 0. A similar analysis yields that the term is 1 if and only if $R_i \geq R_c$.

The second term in the product provides a slight bonus for jobs whose COVID-19 risk rating is lower than the user is willing to handle. This is because a lower COVID-19 risk rating is always beneficial. This multiplier maxes out at 1.12 in the case of $R_i = 4$ and $R_c = 1$, so it is not too influential. Its purpose is mostly to break ties.

$$a_2 = \left(2^{-\max(0, W_i - W_c)}\right)^{E_i + 1}$$

The second term, a_2, is a measure of wages. Each job has a wage bracket W_c associated with it that is an integer from 1 to 4, inclusive. The user also inputs the minimum wage W_i they are willing to work for. We do not consider wage to be an end-all factor for jobs; that is, we assume the user would still be willing to work for a job with lower wages than they would like, albeit not as willingly.

To implement this, we use an exponential decay function. Specifically, if $W_i \leq W_c$ (that is, the job pays at least as much as the user is willing to work for), then this term evaluates to 1 and has no effect on the formula. Otherwise, if the job's wage is less than the user is willing to work for, the function evaluates to $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}$, depending on how many wage brackets W_c and W_i differ.

The outermost exponent comes from E_i, which is 1 if the user values wage over interest and 0 otherwise. Thus, if the user values wage as one of their top priorities, this term is squared to increase its prevalence in the product. Otherwise, nothing happens.

$$a_3 = \left(\frac{1 + 0.25S_i + 0.01D_i}{2.75}\right)^{2 - E_i}$$

The third term, a_3 is the interest term. As defined in Section 3.1, each user has a seven-digit binary string I_i that they input, where the kth bit is 0 if they are not interested in the kth topic and 1 otherwise. Each job also has an associated binary string I_c, defined similarly. We then define S_i as the number of bits that are 1 in both I_i and I_c. In other words, it is the number of interests that the user has that are also in the job. We also define D_i to be the number of bits that are 0 in both I_i and I_c. This is the number of interests that the user does not have that the job does not have as well. Matching interests provide a large bonus, while matching non-interests provide a smaller bonus, mostly for the purpose of breaking ties. This term is 1 in the worst case and 2.75 in the best. We divide by 2.75 to keep this term between 0 and 1 (since all the other terms are). Thus, a job with no matching interests will have a low interest multiplier, while in the best case this multiplier will be 1.

The exponent comes from E_i, a variable which is 1 if the user values wage over interests, and 0 otherwise. If the user does value the interest of the job more than the wage, this term is squared to increase its prevalence in the final product. Otherwise, nothing happens.

$$a_4 = \left(-\frac{\left(\frac{|H_m - H_i + \frac{1}{2}|}{H_m - H_i + \frac{1}{2}}\right)}{2} + \frac{1}{2} \right) \left((0.7)^{H_i - \min(H_M, H_i)} \right)$$

The fourth term, a_4 is a measure of the working hours. Here, H_m and H_M are the minimum and maximum amount of hours one can work at a specific job, respectively, where the scale is from 1 to 3 for low, medium, and high amounts of hours, respectively. The user also inputs H_i, the maximum amount they are willing to work. The first term of the product is constructed similarly to first part of term a_1, so that it is 0 if and only if $H_i < H_m$, and one otherwise. In other words, if the job offers more hours than the user can handle, the term evaluates to 0 and the whole job is discarded.

The second term in the product is an exponential decay function for when $H_i \leq H_m$. If the user is able to work more hours than the job can provide ($H_i > H_M$), this is not a reason to discard the whole job, but it is a reason to lower its overall score. Thus, we use an exponential decay function that lowers the score more the larger the difference between H_M and H_i. This difference is at most 2, so the constant 0.7 was picked since $(0.7)^2 = 0.49$, meaning this term will, at worst, roughly halve the score.

$$a_5 = \frac{\left(\frac{|P_i - P_c + \frac{1}{2}|}{P_i - P_c + \frac{1}{2}}\right)}{2} + \frac{1}{2}$$

The fifth term, a_5, is a measure of physical intensity. Each job has a physical intensity P_c that is 1 to 4, and the user inputs the maximum physical intensity P_i that they are willing to endure. This term is constructed exactly the same as the first part of a_1, meaning it is 0 if and only if $P_i < P_c$ and 1 otherwise.

$$a_6 = \frac{\left(\frac{|A_i - A_c + \frac{1}{2}|}{A_i - A_c + \frac{1}{2}}\right)}{2} + \frac{1}{2}$$

The sixth term, a_6, is a measure of age. Each job has a minimum age A_c, and the user inputs their age A_i. Similar to terms a_5 and the first part of term a_1, this term is 0 exactly when $A_i < A_c$ and 1 otherwise.

Then, the formula is:

$$\prod_{i=1}^{6} a_i$$

For completeness, we write the complete expression for scoring a job:

$$\left(\left(\frac{\left(\frac{|R_i - R_c + \frac{1}{2}|}{R_i - R_c + \frac{1}{2}}\right)}{2} + \frac{1}{2}\right)(1.04)^{R_i - R_c}\right)\left(\left(2^{-\max(0, W_i - W_c)}\right)^{E_i + 1}\right)\left(\left(\frac{1 + 0.25S_i + 0.01D_i}{2.75}\right)^{2 - E_i}\right)\left(\left(-\frac{\left(\frac{|H_m - H_i + \frac{1}{2}|}{H_m - H_i + \frac{1}{2}}\right)}{2} + \frac{1}{2}\right)\left((0.7)^{H_i - \min(H_M, H_i)}\right)\right)\left(\frac{\left(\frac{|P_i - P_c + \frac{1}{2}|}{P_i - P_c + \frac{1}{2}}\right)}{2} + \frac{1}{2}\right)\left(\frac{\left(\frac{|A_i - A_c + \frac{1}{2}|}{A_i - A_c + \frac{1}{2}}\right)}{2} + \frac{1}{2}\right)$$

Note that the score of a job is 0 if and only if any of the following are true, as these are considered to be irreversible conditions:

(1) $R_i < R_c$
(2) $P_i < P_c$
(3) $A_i < A_c$

3.4. **A Mathematical Formula for Internships.**

We now propose a similar mathematical formula to evaluate unpaid internships, which follow a slightly different set of constraints. Specifically, by the assumptions in Section 2.3, we need only consider physical demand, COVID-19 risk, and the interests of the student. Thus, the formula for scoring an internship is the product of three terms instead of six.

We use three terms b_1, b_2, b_3. The first and second terms measure COVID-19 risk and physical demand, respectively. We have:

$$b_1 = a_1, b_2 = a_5$$

Term b_3 is the interests term, but is constructed slightly differently from a_3, the interests term for the job formula. First, we remove E_i, since there is no concept of wages for unpaid internships. We also consider interests to be far more important for internships than jobs; that is, if no interests match up, the interest term should be very close to 0 and the overall score very low. There are 11 interests, so we simply remove the 1 at the beginning of a_3 so that term b_3 maxes out at 1:

$$b_3 = \left(\frac{0.25S_i + 0.01D_i}{2.75}\right)$$

The final formula for scoring an internship is then:

$$\prod_{i=1}^{3} b_i$$

For completeness, we write the complete expression for scoring an internship:

$$\left(\left(\frac{\left(\frac{|R_i - R_c + \frac{1}{2}|}{R_i - R_c + \frac{1}{2}}\right)}{2} + \frac{1}{2}\right)(1.04)^{R_i - R_c}\right)\left(\frac{\left(\frac{|P_i - P_c + \frac{1}{2}|}{P_i - P_c + \frac{1}{2}}\right)}{2} + \frac{1}{2}\right)\left(\frac{0.25S_i + 0.01D_i}{2.75}\right)$$

Once again, this score is 0 if and only if either $R_i < R_c$ or $P_i < P_c$ (or both).

4. Testing the Model

To test our model, we first create a list of 17 high school students, 10 of whom are aiming for paid jobs, and 7 of whom are aiming for internships. We assign each student a list of inputs, all of which are plausible real-life situations. For example, person A's inputs, as seen in Figure 2, can be supported by the following description: Person A is a 17-year-old high school junior who is in need of money and willing to work long hours at any physical level. However, they have many family members who are susceptible to COVID-19 at home, and they are searching for a lower-risk job. They are interested in manual labor, food, and service.

Each of these inputs are plugged into Python code for the respective model. This code implements the model described in Section 3, and can be found in Appendices I and II. Jobs and internships are taken from the data developed as described in Section 3.2. Results of the code (the best jobs and internships for each inputted student), are found in Figures 2 and 3. These matchups are as accurate as they can be with the limited database of jobs. RADAR charts are shown in Figures 4, 5, 6, and 7 for person A and person J. The shaded-in areas are very similar between person A and their selected job as well as person J and their selected job, demonstrating the strength of the model.

Note that the code actually returns the top 10 jobs or internships that would best match with a user. This is done to give the user some choice, and the full output data is shown in Appendices III and IV. However, Figure 2 shows only the best match for succinctness.

Person	R_i	W_i	P_i	H_i	A_i	E_i	I_i	Best Job
A	2	2	4	3	16	1	Food, Service, Manual Labor	Vehicle/Equipment Cleaner
B	4	1	3	1	17	1	Service, Sales	Tutor (remote)
C	4	3	4	3	18	1	Food, Service, Sales, Arts, Technology, Manual Labor, Healthcare	Technology Consultant
D	3	1	3	2	16	0	Service, Arts, Healthcare,	Freelance Musician (remote), Child Care Worker (remote)
E	1	1	1	1	15	0	Food, Service, Arts	Art/Music Education (remote)
F	2	1	3	2	17	0	Sales, Arts, Manual Labor	Freelance Musician (remote), Telemarketer
G	4	3	3	1	18	0	Food, Service, Sales, Technology, Manual Labor, Healthcare	Childcare Worker (in person)
H	4	3	1	1	18	0	Food	Tutor (remote)
I	1	3	2	3	14	0	Food, Service, Sales, Arts, Technology	Telemarketer
J	3	1	2	3	17	1	Service, Sales, Technology	Technology Consultant

FIGURE 2. A table of the 10 job-seeking inputted users and their outputted best jobs. Both jobs are listed in the case of a tie.

Person	R_i	P_i	Interests	Best Internship
K	3	2	Service, Art	Elderly Services (remote)
L	1	1	Communications, Service, Technology	I.T. Support
M	3	3	Life Science, Physical Sciences, Data Analysis, Engineering, Technology	Mechanical Engineering
N	4	3	Education, Life Sciences, Mathematics, Data Analysis	Data Analysis
O	3	2	None	Fund-Raising
P	1	1	Mathematics, Data Analysis, Finances	Marketing
Q	4	4	Service	Elderly Services (remote)

FIGURE 3. A table of the 7 internship-seeking inputted users and their outputted best internships. Both internships are listed in the case of a tie.

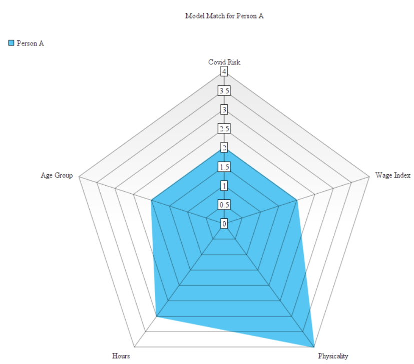


FIGURE 4. A RADAR chart depicting the inputs of Person A.

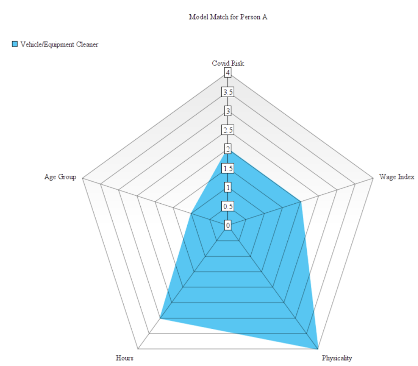


FIGURE 5. A RADAR chart depicting the attributes of Person A's outputted best job.

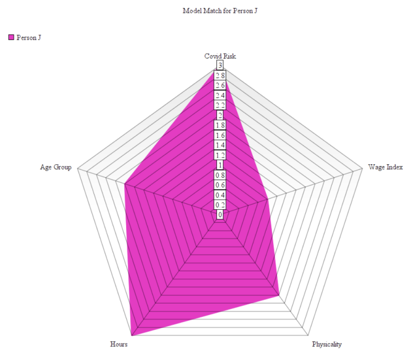


Figure 6. A RADAR chart depicting the inputs of Person J.

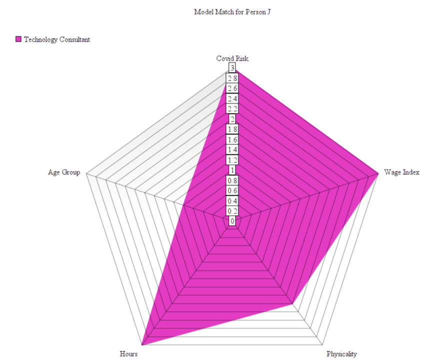


Figure 7. A RADAR chart depicting the attributes of Person J's outputted best job.

4.1. **Sensitivity Analysis.**

To test the sensitivity of our model, we can examine Person B. Because the maximum hours for Person B is level 1, their jobs are limited to those with a time commitment of level 1. No jobs related to service or sales had a small enough time commitment, so Person B was unable to get any recommendations with jobs matching his interests. However, if we change $H_1 = 1$ and $H_1 = 2$, we find that the model is able to recommend Sales and Service related jobs like News/Street Vendor and Technology Consultant as top choices. Term a_4, as in Section 3.3, lowers the score if the maximum hours offered by a job is less than the user's inputted hours. Thus, preference was given to jobs that could support level 2 hours (part-time work). Moreover, an increase in the amount of maximum hours increases the pool of available jobs, making jobs which align with Person B's interest score higher.

Another factor we can change slightly to get different results is the wage level. Because the wage Person D chose was $W_i = 1$, wage did not affect the comparisons between jobs. Thus, the top jobs were calculated mainly to fit his their interests and safety, like Freelance Musician (remote) and Childcare Worker (remote). However, if we modify the wage choice from $W_i = 1$ to $W_i = 2$, we find that these low paying jobs are replaced with Animal Shelter Worker and Freelance Musician (in person). Term a_2, as in Section 3.3, halves the score if the wage level of the job is 1 less than the wage level the user inputs. Thus, the scores of the two previous top jobs were significantly decreased and the two new jobs with higher pay were given a better ranking.

5. Presenting the Model

The models are presented using a website. Our model requires user input in the form of answers to many questions, so we decided a questionnaire was how we would gather user data. This questionnaire can be best represented through a website. We also need to do calculations in the background, so physical options like newspapers do not suffice.

The first question, on the homepage, determines whether or not the user is looking for a paid job or an unpaid internship. Since the models for the two are different, this question allows the computer to know which model to use. Each answer leads to a separate form. If the user enters that they are looking for a job, they then answer the following questions on the respective form, each of which is accompanied by a set of choices with labels, as shown in Figure 8:

(1) What is your age, in years?
(2) Regarding COVID-19 risk, what is the maximum level of contact of your desired job?
(3) What is your preferred minimum hourly wage?
(4) What is the maximum daily hours you can work?
(5) What is the maximum level of physical exertion you would prefer on the job?
(6) Please check off your fields of interest!
    (a) Food
    (b) Service
    (c) Sales
    (d) Art
    (e) Technology
    (f) Manual Labor
    (g) Healthcare

If the user desires an internship, they fill out a shorter form:

(1) Regarding COVID-19 risk, what is the maximum level of contact of your desired job?
(2) What is the maximum level of physical exertion you would prefer on the job?
(3) Please check off your fields of interest!
    (a) Education
    (b) Life Sciences
    (c) Physical Sciences
    (d) Mathematics
    (e) Data Analysis
    (f) Engineering
    (g) Finances
    (h) Communications
    (i) Service
    (j) Art
    (k) Technology

After the user fills out their respective form, the computer performs the calculations based on the correct model. It is assumed there is access to a database of jobs near the user's location. Based on their inputs, the computer computes a score for each job in proximity to the user and outputs the ten best jobs, accompanied by their scores. The top ten jobs are printed out as to give the user some choice, because all of the top ten jobs would ideally be suitable fits.

Figure 8 depicts a prototype of what the website would look like.

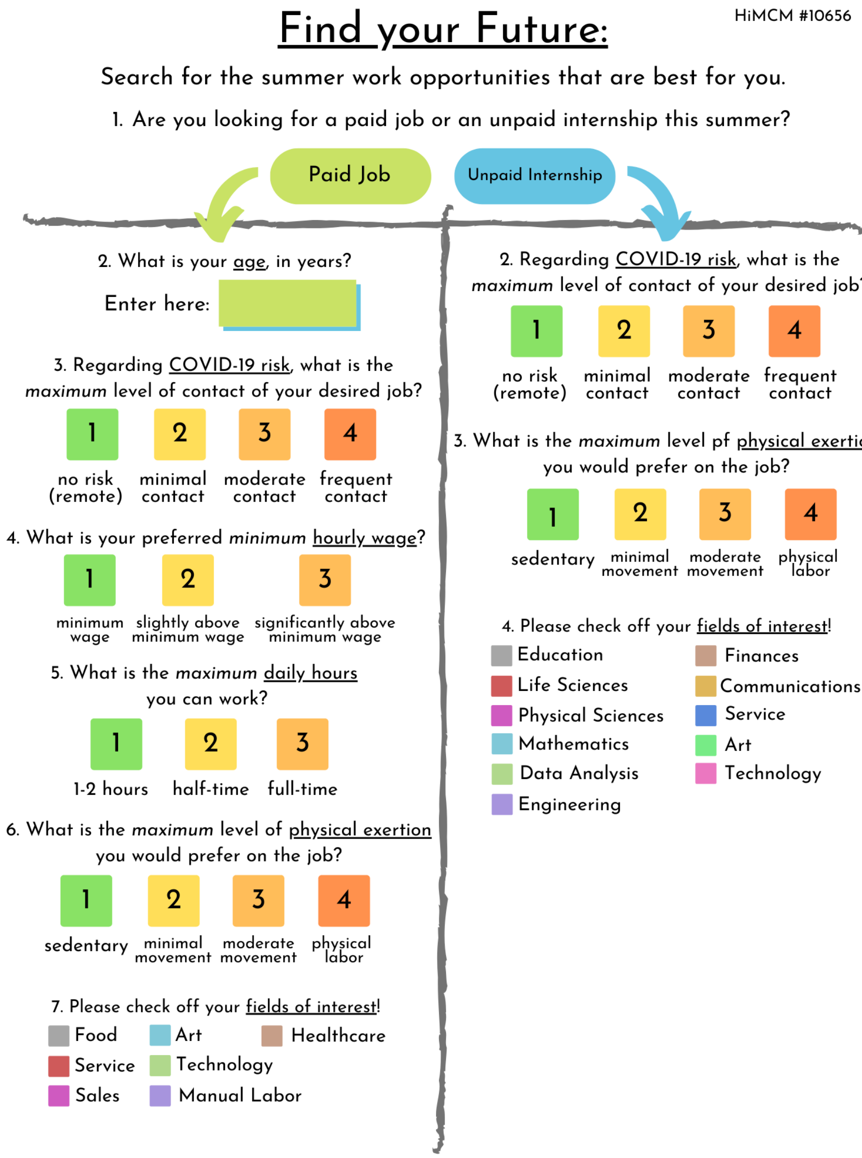


Figure 8. The setup of the website.

6. Strengths and Weaknesses

Our models compare different jobs and internships based on user input very well, but there are still improvements to be made. We list some strengths and weaknesses of the model in Figure 9:

Fields	Strengths	Weaknesses
Interpretation of Problem	We consider unpaid internships and jobs separately, allowing for maximum control over choice of summer occupation. Both types of work are fitted to the user's preferences and demographics, allowing for a job of best fit. We consider the possible motivations for choosing a job, which we determined to be either interest-driven or pay-driven.	Many 18-year olds hold high school diplomas, which allow for more job opportunities. While many require extra training afterwards, they are still potential jobs that were not considered. We assume that work schedules were consistent; i.e. the same number of hours was worked every day. Of course, this is not the case for the vast majority of jobs, which is a weakness of our model and something that can be addressed in the future.
Collection of Data	Job data is gathered on a federal level, which ensures a uniform standard for all job opportunities available to our selected population. Job data is pruned to consider jobs which require practically no qualifications, which would be appropriate for high school applicants. General topics of interest surrounding each job are chosen based on historical data about the most popular fields for high school students during the summer. This ensures that we have a strong and realistic baseline for our final list of jobs.	Job descriptions are unable to quantify physical demand, so many assumptions were made. The same is true for hours and wages. Commute is not considered as a potential factor. A differentiation is made between remote and in-person jobs (a job is remote if and only if its COVID-19 risk rating is 1), but no distinction is made between short and long commutes, as this would force our test data to be extremely specific to location. This would also involve a large amount of user and job data matching.
Formation of Model	Both formulas are componentized, allowing each subpart to be easily adjusted mathematically. In our model, the importance of each factor/variable is taken into account and is reflected in the operations and constants of the formula. Each term in the final multiplication of terms is adequately justified. The model outputs a list of the top 10 scored jobs it determines fit for the user, allowing the model to be flexible. Consider a user whose top ranked job is determined to be unfit for them. Despite this inconvenience, the user can choose the next best job and still be assured that it is right for them.	Constants in the formula are somewhat arbitrary, although they do reflect the intended behavior of the respective term. Job scores are consistent relative to other jobs for the same applicant but are not consistent across applicants. That is, a 1.6 for one person could still be a worse fit than a 1.2 for the other person, but within one applicant's data, relative positioning is accurate.
Testing of Model	Our test cases are realistically justified and test the extreme and moderate cases covered by the model. Our test cases are not randomly generated or assigned, but rather carefully selected to provide a diverse array of potential applicant profiles.	Our testing data does not cover all types of people under variables given by the model. However, the model should still work for all real-life users which fit the initial requirements.

FIGURE 9. A table depicting the strengths and weaknesses of our model.

7. Conclusion

In this paper, we developed a model to find the best work opportunity for a high school student in the summer of 2021, whether that be a job or unpaid internship. This model is based off the user's input of a variety of relevant factors, such as COVID-19 risk, interests, and desired wage (in the case of a job). Each job and internship has assigned statistics for the same factors, and these are then combined to score each work opportunity with respect to the user. Two different models were created for jobs and unpaid internships, because they operate off a slightly different set of factors. However, the models are very similar in operation and sensitivity.

The model was tested with 17 different data inputs in the form of fictional students, with 10 students seeking jobs and 7 seeking internships. Jobs were selected for each of the 10 users from a database of 36 jobs, as curated in Section 3.2. A similar list of 25 internships was developed for the 7 users seeking those opportunities. The input data was designed to be as realistic as possible while covering a large gamut of possible inputs to the model. The model was found to work sufficiently in all input cases, providing reasonable work opportunities for all users, as desired. These opportunities matched the user's preferences as well as possible from the small database available. We posit that, with a larger database of potential work opportunities, matches will only become more accurate.

The model was presented using a website for the greatest ease of entering user input as well as the fact that all calculations could be run automatically. In the future, we desire to fully implement this model. We also hope to improve the model by factoring in commute and favoring jobs that are closer to home over jobs that are farther away. While it was infeasible to manually add commute data for each job on each input case, this could be done in a real implementation via a web scraper that scrapes job data from online using the user's location. Another factor worth considering is qualifications. These were disregarded in the current model via Assumption 4, but the inclusion of these would make the model more accurate and realistic.

References

[1] Find occupations. [Online]. Available from: https://www.onetonline.org/find/. O*NET Online.

[2] Child labor provisions of the fair labor standards act (flsa) for nonagricultural occupations. [Online]. Available from: https://www.dol.gov/agencies/whd/fact-sheets/43-child-labor-non-agriculture, December 2016. U.S. Department of Labor.

[3] Summer employment: A snapshot of teen workers : Career outlook. [Online]. Available from: https://www.bls.gov/careeroutlook/2017/article/youth-summer-employment.htm, June 2017. U.S. Bureau of Labor Statistics.

[4] Top 20 fields for internships: Get your skills on. [Online]. Available from: https://www.nytimes.com/2017/02/03/education/edlife/top-20-fields-for-internships-get-your-skills-on.html?auth=login-google, February 2017. The New York Times.

[5] Discover your career path. [Online]. Available from: https://www.owlguru.com/career/path/, March 2020. OwlGuru.

[6] Guidance on preparing workplaces for covid-19. [Online]. Available from: https://www.osha.gov/Publications/OSHA3990.pdf, 2020. Occupation Safety and Health Administration.

[7] Home : Occupational outlook handbook. [Online]. Available from: https://www.bls.gov/ooh/, September 2020. U.S. Bureau of Labor Statistics.

[8] Kimberly Railey. Early high school graduation programs gain traction. [Online]. Available from: https://www.usatoday.com/story/news/nation/2013/07/27/early-high-school-graduation/2518817/, July 2013. USA Today.

8. Appendix I: Python Code for the Job Model

```
import csv

people_file = open('Job Database - JobPeople.csv', 'r')
peopl = csv.reader(people_file)
people = [line for line in peopl]
people.pop(0)

jobs_file = open('Job Database - Jobs.csv', 'r')
jo = csv.reader(jobs_file)
jobs = [line for line in jo]
jobs.pop(0)

def partition(arr, low, high):
    i = (low - 1)
    pivot = arr[high][1]

    for j in range(low, high):
```

```
        if arr[j][1] <= pivot:
            i = i + 1
            arr[i], arr[j] = arr[j], arr[i]

    arr[i + 1], arr[high] = arr[high], arr[i + 1]
    return (i + 1)

def quickSort(arr, low, high):
    if len(arr) == 1:
        return arr
    if low < high:
        pi = partition(arr, low, high)

        quickSort(arr, low, pi - 1)
        quickSort(arr, pi + 1, high)

def score(R,r,W,w,Inte,inte,P,p,H,hm,hM,A,a,poe):
    rScore = (abs(R-r+0.5)/(R-r+0.5)/2.0+0.5)*(1.04**(R-r))
    wScore = 2.0**(-max(0,W-w))
    count0 = 0
    count1 = 0
    for i in range(7):
        if Inte[i] == inte[i]:
            if Inte[i] == '0':
                count0 += 1
            if Inte[i] == '1':
                count1 += 1
    iScore = (1+(0.25*count1) + (0.01*count0))/2.75
    pScore = abs(P-p+0.5)/(P-p+0.5)/2.0+0.5
    hScore = (-0.5)*(abs(hm-H-0.5)/(hm-H-0.5)-1)*(0.7**(H-min(hM,H)))
    aScore =  abs(A-a+0.5)/(A-a+0.5)/2.0+0.5
    return rScore * (wScore ** (poe+1)) * (iScore ** (2-poe)) * pScore *
           hScore * aScore

recs = []

for person in people:
    scores = []
    for job in jobs:
        scores.append([job[0], score(int(person[1]), int(job[2]),
                       int(person[2]), int(job[2]), person[3], job[7],
                       int(person[4]), int(job[4]), int(person[5]),
                       int(job[5]), int(job[6]), int(person[6]),
                       int(job[2]), int(person[7]))])
    quickSort(scores, 0, len(scores)-1)
```

```
    top10 = scores[::-1][0:10]
    print(scores[::-1][0:10])
    recs.append([person[0], ''])
    for rec in top10:
        if rec[1] > 0:
            recs.append(rec)

recs_file = open('recs_file.csv', 'w')
csvwriter = csv.writer(recs_file)
csvwriter.writerows(recs)
```

9. Appendix II: Python Code for the Internship Model

```
import csv

people_file = open('Job Database - InternPeople.csv', 'r')
peopl = csv.reader(people_file)
people = [line for line in peopl]
people.pop(0)

interns_file = open('Job Database - Internships.csv', 'r')
inter = csv.reader(interns_file)
interns = [line for line in inter]
interns.pop(0)

def partition(arr, low, high):
    i = (low - 1)
    pivot = arr[high][1]

    for j in range(low, high):
        if arr[j][1] <= pivot:
            i = i + 1
            arr[i], arr[j] = arr[j], arr[i]

    arr[i + 1], arr[high] = arr[high], arr[i + 1]
    return (i + 1)


def quickSort(arr, low, high):
    if len(arr) == 1:
        return arr
    if low < high:
        pi = partition(arr, low, high)

        quickSort(arr, low, pi - 1)
        quickSort(arr, pi + 1, high)
```

```
def score(R,r,Inte,inte,P,p):
    rScore = (abs(R-r+0.5)/(R-r+0.5)/2.0+0.5)*(1.04**(R-r))
    count0 = 0
    count1 = 0
    for i in range(11):
        if Inte[i] == inte[i]:
            if Inte[i] == '0':
                count0 += 1
            if Inte[i] == '1':
                count1 += 1
    iScore = (0.25*count1 + 0.01*count0)/2.75
    pScore = abs(P-p+0.5)/(P-p+0.5)/2.0+0.5
    #print(rScore, iScore, pScore)
    return rScore*iScore*pScore

recs = []

for person in people:
    scores = []
    for intern in interns:
        scores.append([intern[0], score(int(person[1]),
                       int(intern[2]), person[3], intern[1],
                       int(person[2]), int(intern[3]))])
    quickSort(scores, 0, len(scores)-1)
    top10 = scores[::-1][0:10]
    print(scores[::-1][0:10])
    recs.append([person[0], ''])
    for rec in top10:
        if rec[1] > 0:
            recs.append(rec)

irecs_file = open('irecs_file.csv', 'w')
csvwriter = csv.writer(irecs_file)
csvwriter.writerows(recs)
```

10. Appendix III: Full Job Recommendations for All Test Cases

A	**Scores**	**F**	**Scores**
Vehicle/Equipment Cleaner	0.56	Freelance musician (virtual)	0.225314
Municipal Services	0.469091	Telemarketer	0.225314
Ticket Taker	0.469091	Freelance painter	0.220046
Garbage Collector	0.328364	Garbage Collector	0.220046
Parking assistant	0.328364	Janitor	0.220046
Janitor	0.328364	Theatre Broadcast technician	0.220046
Freelance painter	0.262182	Art/Music Education (remote)	0.15772
Theatre Broadcast technician	0.262182	Childcare Worker (virtual)	0.143076
Childcare Worker (virtual)	0.121018	Parking assistant	0.140284
Telemarketer	0.096436	Ticket Taker	0.140284
B	**Scores**	**G**	**Scores**
Tutor (remote)	0.531754	Childcare Worker (in person)	0.218328
Art/Music Education (remote)	0.527663	Freelance musician (in person)	0.214876
Childcare Worker (virtual)	0.527663	Janitor	0.11353
Freelance musician (virtual)	0.527663	Childcare Worker (virtual)	0.084787
Freelance musician (in person)	0.487855	Theatre Broadcast technician	0.071511
Janitor	0.409041	Freelance painter	0.071511
Theatre Broadcast technician	0.409041	Tutor (remote)	0.059036
Freelance painter	0.409041	Art/Music Education (remote)	0.058102
Childcare Worker (in person)	0.393309	Freelance musician (virtual)	0.058102
C	**Scores**	**H**	**Scores**
Technology Consultant	0.661818	Tutor (remote)	0.040997
Oil/Gas Service Unit Operator	0.567273	Art/Music Education (remote)	0.04022
Fast Food Counter Worker	0.567273	Childcare Worker (virtual)	0.04022
Order Filler	0.567273	**I**	**Scores**
Electrician Helper	0.567273	Telemarketer	0.076377
Telecom Equipment Technician	0.567273	Freelance musician (virtual)	0.053464
Veterinary Receptionist	0.567273	Childcare Worker (virtual)	0.052483
Animal Shelter Worker	0.567273	Art/Music Education (remote)	0.037425
Retail Store Cashier	0.545455	Tutor (remote)	0.026126
Bartender	0.545455	**J**	**Scores**
D	**Scores**	Technology Consultant	0.650909
Freelance musician (virtual)	0.33919	Telemarketer	0.605696
Childcare Worker (virtual)	0.33919	Telecom Equipment Technician	0.56
Freelance musician (in person)	0.3136	Childcare Worker (virtual)	0.503436
Animal Shelter Worker	0.3136	Ticket Taker	0.487855
Art/Music Education (remote)	0.237433	Fast Food Counter Worker	0.465455
Theatre Broadcast technician	0.228848	Veterinary Receptionist	0.465455
Ticket Taker	0.228848	Order Filler	0.465455
Parking assistant	0.228848	News/Street Vendor	0.392
Freelance painter	0.228848	Freelance musician (virtual)	0.352405
Childcare Worker (in person)	0.220046		
E	**Scores**		
Art/Music Education (remote)	0.3136		
Tutor (remote)	0.220046		
Childcare Worker (virtual)	0.216648		

FIGURE 10. A table depicting the model's job recommendations for all test cases.

11. Appendix IV: Full Internship Recommendations for All Test Cases

K	Scores	O	Scores
Elderly Services (remote)	0.133725	Fund-Raising	0.039331
Customer Service (remote)	0.129792	Editing	0.039331
I.T.Support	0.129792	Mathematics Research (Remote)	0.039331
Graphic Design (remote)	0.129792	Journalism	0.039331
Public health and Safety	0.116364	Elderly Services (remote)	0.039331
Mathematics Research (Remote)	0.031465	Event Planning	0.039331
Journalism	0.031465	Customer Service (remote)	0.035398
Event Planning	0.031465	I.T.Support	0.035398
Fund-Raising	0.031465	Graphic Design (remote)	0.035398
Editing	0.031465	software development	0.035398
L	**Scores**	**P**	**Scores**
I.T.Support	0.210909	Marketing	0.298182
Customer Service (remote)	0.210909	Data analysis	0.210909
PR/Social Media	0.210909	Mathematics Research (Remote)	0.12
Fund-Raising	0.12	software development	0.116364
Event Planning	0.12	Editing	0.025455
Journalism	0.12	Event Planning	0.025455
Elderly Services (remote)	0.12	Elderly Services (remote)	0.025455
Editing	0.12	Fund-Raising	0.025455
Graphic Design (remote)	0.116364	Journalism	0.025455
software development	0.116364	Customer Service (remote)	0.021818
M	**Scores**	**Q**	**Scores**
Mechanical Engineering	0.381818	Elderly Services (remote)	0.143165
Lab assistant	0.290909	I.T.Support	0.139074
Data analysis	0.21632	Customer Service (remote)	0.139074
software development	0.21632	Police Work	0.132364
Marketing	0.208454	Automobile Service	0.132364
Mathematics Research (Remote)	0.117993	Elderly Services (in person)	0.127273
Graphic Design (remote)	0.11406	Public health and Safety	0.1248
PR/Social Media	0.11406	Patient Care	0.123636
I.T.Support	0.11406	Customer Service (in person)	0.123636
Public health and Safety	0.101818	Life Science Museum Guide	0.116364
N	**Scores**		
Data analysis	0.233154		
Marketing	0.224973		
Lab assistant	0.211782		
Life Science Museum Guide	0.2		
Mathematics Research (Remote)	0.130893		
software development	0.126803		
Teaching	0.116364		
Public health and Safety	0.113455		
Patient Care	0.112727		
Mechanical Engineering	0.109673		

FIGURE 11. A table depicting the model's internship recommendations for all test cases.

12. APPENDIX V: FULL JOB AND TEST DATABASE

Jobs	Age Bracket	COVID Risk	Wage	Phys Demand	minHours	maxHours	Interests
Restaurant Cook	14	3	2	3	2	2	1000000
Fast Food Cook	14	3	1	3	2	2	1000000
Busser	16	3	1	3	2	2	1000000
Waiter	16	4	1	3	2	2	1100000
Order Filler	14	3	1	2	2	3	1100000
Fast Food Counter Worker	14	3	1	2	2	3	1100000
Grocery Store Cashier / Bagger	14	4	1	2	2	3	0100000
Retail Store Cashier	14	4	1	2	2	3	0110000
Ticket Taker	14	2	1	1	2	3	0100000
News/Street Vendor	16	3	2	2	2	2	0110000
Technology Consultant	14	3	3	2	2	3	0110100
Telecom Equipment Technician	14	3	3	2	2	3	0100100
Telemarketer	14	1	2	1	2	3	0010100
Freelance musician (remote)	14	1	3	2	1	2	0101000
Freelance musician (in person)	14	3	3	2	1	2	0101000
Freelance painter	14	2	3	2	1	2	0001000
Theatre Broadcast technician	14	2	3	2	1	2	0001000
Material Moving Worker	16	3	2	3	2	3	0000010
Assembly Line Manufacturing	16	3	2	3	3	3	0000010
Vehicle/Equipment Cleaner	14	2	2	4	1	3	0100010
Carpenter Helper	16	3	2	3	2	3	0000010
Electrician Helper	16	3	3	3	2	3	0000110
Farmworker	14	3	2	4	3	3	0000010
Garbage Collector	14	2	3	3	2	2	0000010
Municipal Services	16	2	3	4	2	3	0000010
Bartender	18	4	1	3	2	3	0110000
Hoist and Winch Operator	18	3	3	4	3	3	0000010
Oil/Gas Service Unit Operator	18	3	3	4	3	3	0100001
Animal Shelter Worker	14	3	2	3	2	3	0100001
Veterinary Receptionist	16	3	2	2	3	3	0100001
Childcare Worker (remote)	14	1	2	1	1	3	0100001
Childcare Worker (in person)	14	3	1	3	1	3	0000001
Parking assistant	14	2	2	2	2	2	0100000
Janitor	16	2	2	3	1	2	0000010
Tutor (remote)	14	1	2	1	1	1	0100000
Art/Music Education (remote)	14	1	2	1	1	1	0101000

Job Interest Legend

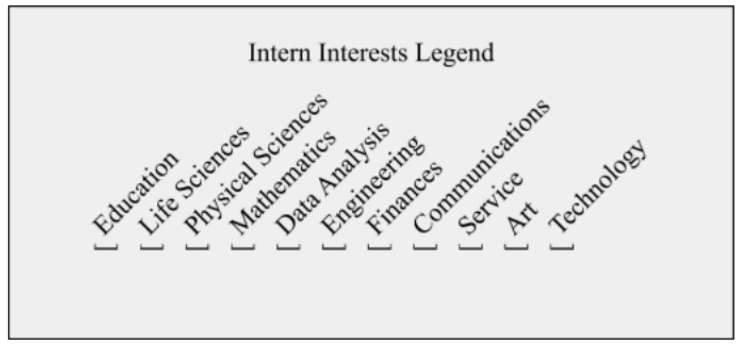
Food, Service, Sales, The Arts, Technology, Manual Labor, Healthcare

Person	COVID Risk	Wage	Interests	Phys Demand	Hours	Age	Experience(0) / Pay(1)
A	2	2	1100010	4	3	16	1
B	4	1	0110000	3	1	17	1
C	4	3	1111111	4	3	18	1
D	3	1	0101001	3	2	16	0
E	1	1	1101000	1	1	15	0
F	2	1	0011010	3	2	17	0
G	4	3	1110111	3	1	18	0
H	4	3	1000000	1	1	18	0
I	1	3	1111100	2	3	14	0
J	3	1	0110100	2	3	17	1

FIGURE 12. The jobs and test cases that our model was tested with.

13. Appendix VI: Full internship and test database

Internships	Interests	COVID Safety	Phys Demand
Customer Service (in person)	00000001100	4	1
Customer Service (remote)	00000001100	1	1
Marketing	00011011000	1	1
Mechanical Engineering	00110100001	3	3
Journalism	00000001000	1	1
Editing	00000001000	1	1
Data analysis	00011000000	1	1
software development	00001000001	1	1
Graphic Design (remote)	00000000011	1	1
Lab assistant	01110000000	3	3
Patient Care	01000000100	4	3
Mathematics Research (Remote)	00010000000	1	1
Public health and Safety	01000001100	3	2
Fund-Raising	00000001000	1	1
Event Planning	00000001000	1	1
Teaching	10000000000	4	3
I.T.Support	00000000101	1	1
Life Science Museum Guide	11000001100	4	3
PR/Social Media	00000001001	1	1
Elderly Services (in person)	00000000100	4	3
Elderly Services (remote)	00000000100	1	1
Photographer	00000000010	3	3
Automobile Service	00000000100	3	4
Political Campaign Assistant	00000001000	3	3
Police Work	00000000100	3	3



Person	COVID Risk	Phys Demand	Interests
K	3	2	00000000110
L	1	1	00000001101
M	3	3	01111100001
N	4	3	11011000000
O	3	2	00000000000
P	1	1	00011010000
Q	4	4	00000000100

FIGURE 13. The internships and test cases that our model was tested with.