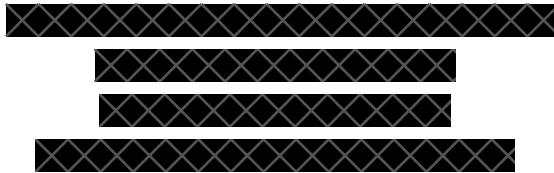


Research & Innovation (CS7CS6)

██████████ Academic Track Prediction Algorithm



1 Academic Track Prediction Algorithm: A Preview

A machine learning algorithm takes in a researcher's citation count, number of publications, h-index and past employing universities of the past 10 years in order to predict these statistics 5 years into the future, including a list of likely future employers. This is the centerpiece of our assigned dissertation (O'Driscoll, 2018), and the main product we would offer to paying researchers and universities through a well-advertised website. We believe we stand to help underappreciated, high-potential researchers accelerate their careers and prevent current well-performing individuals from slacking off. Moreover, we are patching up the main issue with the algorithm that the dissertation's author cites - its small sample size - by constantly incorporating the users of the product as new points of data for improved predictions.

2 Business Model Canvas

Summary information on all relevant features of the business model are categorically displayed in figure 1 on the next page. The remainder of this section provides additional notes and information for the important categories requiring it.

At the core of the value proposition is the idea that our product helps researchers understand their potential and capabilities, and the unique underlying algorithm allows researchers to choose a suitable university and thus pick a better future direction. Researchers need to consider their reputation, income, and the resources needed for their craft, giving the paid usage of the algorithm a wide group of potential customers. However, because the concept of our product is relatively new and may not be widely accepted immediately, we have introduced a free, reduced-information trial option besides the complete paid version for individual researchers. This trial

would only offer insight into 5 suitable universities, nudging the indecisive forward while creating dialogue about the product as all researchers could in theory discuss surprisingly good (or bad) predictions. For universities, our product can be widely used as a factor in hiring decision, to evaluate the status of our researchers, and to help plan their investments. Therefore, we will consider setting a higher price for the subscription model addressed to universities.

Since the primary role of our product is forecasting, we believe that our primary researcher customer base is three groups of people who are at a turning point in their careers. The first is people who are at the top of their careers but whose research level will decline, the second is people whose current research level is low but who will improve greatly in the future, and the third is people who are recent PhD graduates who are confused about their future. All three types of people are likely to need our products for forecasting, for selecting their target universities and for planning their future. Our university customers may have an easier reason to use our product, as every university wants to hire more promising and strong researchers, especially for the less prestigious universities, and our product is a good opportunity for them to select promising researchers and improve their reputation in the future.

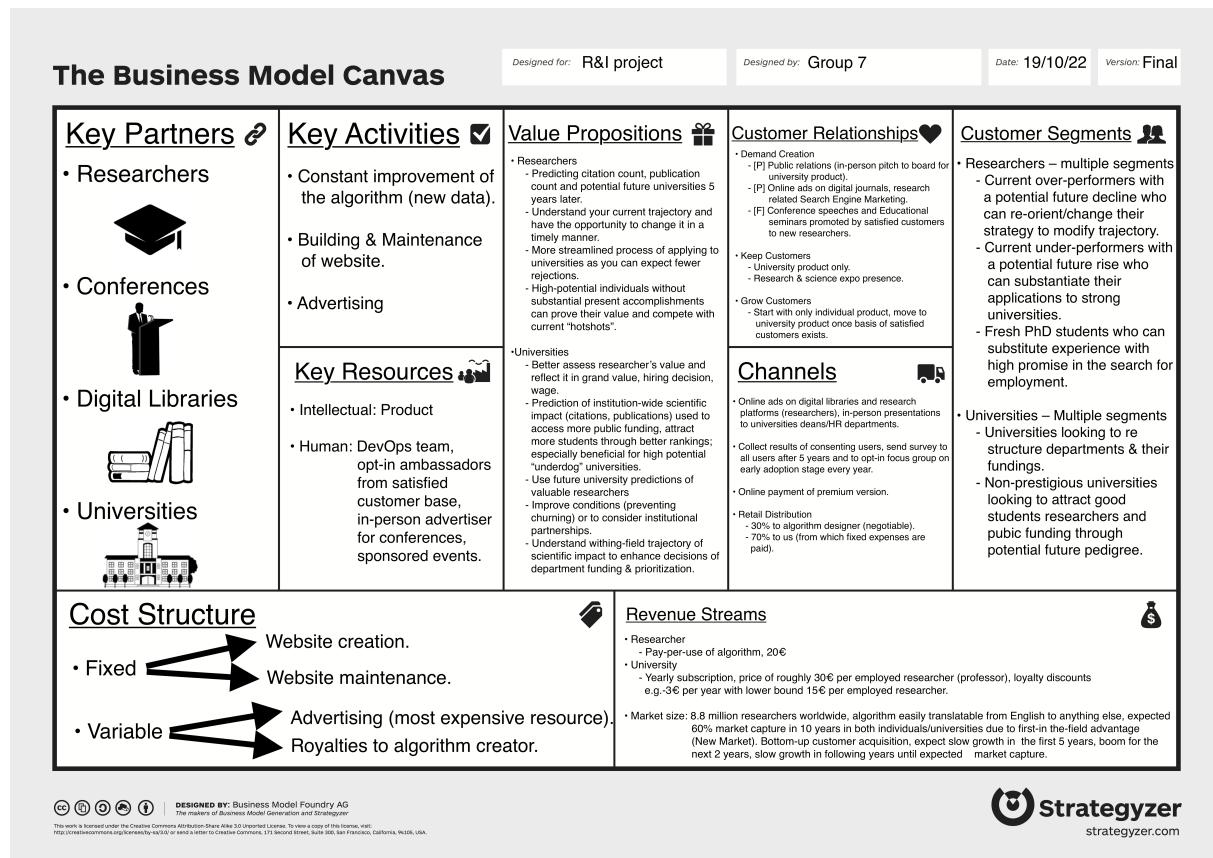


Figure 1: Business Model Canvas, including summary information on the product's value proposition, customer segments, etc.

As far as channels are concerned, the main way to promote our product is through advertising, and we plan to promote it in e-libraries and research platforms, as our target customers are researchers who will frequently visit academic-related websites, helping to increase the visibility of our product. The algorithm itself is accessed only through the Internet, eliminating a significant portion of costs. To ensure that our customers can use our products with confidence, we also provide a range of after-sales services and collect feedback from our users. This is to improve our reputation and to increase user stickiness. For users whose predictions turn out significantly off base - outside a pre-determined confidence interval, for instance - refunds will be provided. Finally, we share 30% of our revenue with the authors of the algorithms, as they are a core part of the product.

In order for us to maintain a good relationship with our customers, we will take several approaches, for example, conference speeches and educational seminars, and in our plan, as our product becomes more known among users, we hope to promote it to more users in this way. In addition to gaining new users, we also want to keep our old customers, and we will contact them regularly or conduct research- & science-expo events. We carefully considered these aspects of customer relationships when assessing revenue streams, concluding that as researchers are normally upper-middle class or above and universities can convert powerful research into financial value, the paying capacity of our customers is quite satisfactory.

3 Formulation and Testing of Hypotheses

We have identified 6 core premises on which the product concept - in its initial form - hinges to succeed. As we intended to begin by launching only the researcher's version, the premises mostly focus on its success factors in isolation of the universities' version, except for the final one.

1. Researchers find the information provided by the algorithm useful.
2. Researchers are willing to pay around 20 euro to find them.
3. Researchers are more inclined to use the algorithm if others use it.
4. Researchers prefer confidentiality for their results.
5. Researchers do not find the product detrimental to research quality.
6. Researchers would not avoid universities that use the algorithm.

To get a grasp of how potential customers would receive the product, we conducted interviews with a handful of our friends and acquaintances that intend to pursue a career in academic research. Our questions closely followed the 6 aforementioned premises, inquiring for instance about willingness to pay, perceived peer pressure to use the product, or specific features of the

product that interested the interviewees. These interviews helped us gain confidence in the project's utility for undecided researchers and people with impostor syndromes, as well as the project's motivational value for underperforming and underappreciated researchers. However, we were happy to receive constructive criticism and consequently modify the product concept as well, in ways described in the remaining paragraphs of this section.

There was universal consensus on the willingness to use both the citation/publication predictor and the university suggestion element of the product and pay at least 10 euro for it. Only a third of the offered sums exceeded 20 euro though, prompting our first pivot to **decrease the price of the premium package to 15 euro**. Responses to peer pressure are split, as only half of our respondents declared they would be affected by others using the algorithm. We would consequently pivot to **expecting the boom phase of customer acquisition to be delayed by one or two years** over the BMC plan.

Our assumption of preferences towards confidential results was unanimously confirmed, unlike the hypothesis that researchers would not avoid algorithm-using universities. While some respondents declared themselves unconcerned as universities already use some non-algorithmic quotas and the product's influence would not weigh too much in the universities' relations with their researchers, more than half said they would avoid the universities using the algorithm or reconsider going into research altogether. The most often cited reason is fear that a researcher will be fired or denied employment based on the algorithm's results, which prompted us to restrict universities' product usage to their current employees; we can achieve this by **maintaining current employee lists from subscribed universities and applying the algorithm following specific requests rather than offering discretionary access to it**.

Lastly, opinions were also split on the impact on research quality. Even though most respondents confirmed our hypothesis affirming they did not believe there would be a negative impact, concerns were raised that underprivileged researchers would be further discouraged from going into research, as the regrettable setbacks suffered by past underrepresented groups would be reflected in the predictions for the new generation of researchers. While this did not necessarily prompt a pivot, it made us aware of the importance of **only basing the algorithm on the researchers' years in the field and the relevant scientific prominence metrics**, ignoring other personal traits which should not impact the researchers' opportunities and career in the first place.

4 Ethics Canvas

Figure 2 is a reflection of our ethical considerations of the algorithm and its potential impact on researchers, universities and the overall process of research alike. If the product's predictive power is overestimated and outweighed in employment decisions, unfair hiring practices, shattered self-esteem and poor quality research born of algorithm gamification may abound. It is for this reason that we endorse the dissertation author's claim towards the end of their work, that the results of the algorithm ought to always be considered holistically alongside a researcher's other strengths and aspirations, and that the human intuition must not be divorced from the computational suggestion. As long as the algorithm is seen as a tool of information rather than an inherent decision maker and we impose the necessary checks on universities to prevent abuse in the hiring process, we believe the product can do significantly more good than harm in the world overall.

Academic Track Prediction Algorithm				
Ethics Canvas, Group: 7 Date: 21/10/2022				
Individuals Affected	Behaviour	What can we do?	Worldviews	Groups affected
<ul style="list-style-type: none"> Researchers PHD students 	<ul style="list-style-type: none"> Researchers prefer to publish better quality papers Less researcher willing to do research that takes a long time. Researchers are more likely to change universities based on algorithm result. Universities are more likely to fire researchers. 	<ul style="list-style-type: none"> Add more aspects to evaluate or predict the ability of researchers. This forecast data is confidential and can only be viewed by the researcher himself and the employer. Disclaimer about avoiding the exclusive use of the algorithm in employment decisions and considering other relevant, human-determined criteria (team chemistry, loyalty, character etc.) 	<ul style="list-style-type: none"> Easier to find higher positions for competent researchers. Universities are more likely to identify underachieved researchers. Researchers will face more pressure. 	<ul style="list-style-type: none"> Universities Research institute Digital libraries Conference
Relations				Group Conflicts
<ul style="list-style-type: none"> There may be discrimination among colleagues because of underwhelming algorithm results 				<ul style="list-style-type: none"> Researchers are more likely to change jobs, so competition between universities is more intense. Underperforming researchers may be more likely to lose their job
Product or Service Failure		Problematic Use of Resources		
<ul style="list-style-type: none"> Prediction inaccuracy may unfairly diminish a researcher's self-esteem or job prospects Researchers with few published articles inherently disadvantaged in terms of accuracy 		<ul style="list-style-type: none"> more data centers are needed to store data about researchers important to maintain confidentiality of users' personal data (names, prediction results) to protect employment prospects 		

Figure 2: Ethics Canvas, including summary information on the product's potential to negatively affect groups or individuals and measures we can take to suppress it.

5 Afterword - Work Distribution

Our team consists of four people all working equally hard towards implementing the Research and Innovation Methods group project. As a team we have met four times. In the first meeting we took our time in to get to know each other and try to understand the project. First, we went through the project instructions and figured out what our initial steps should be. Then we read the paper assigned to us in order to develop a deep understanding of it. In addition, we took notice of some key parts of the paper to help us later; in case we need to revisit the paper and extract useful information. Next, we discussed and developed a theoretical model of how we could make the paper's product work as a business.

Then after gathering more thoughts on the subject, we met again a week later to exchange our ideas. Also, we got into work and started creating the BMC (Business Model Canvas). During the creation of the BMC, we went back and forth with the lecture notes in order to get insights on what we must include and where. The value proposition, customer segments and customer relationships fields were completed by the end. The third meeting saw the completion of the BMC, as well as the conception of a survey which we later handed out to our future researcher colleagues for the product testing and pivoting task. All of us shared great ideas and concepts. At the end of the meeting, we split equally among us the workload that had to be done. So, one person was responsible for creating the Ethics Canvas, another one was responsible for making the BMC clean and tidy while also composing the work organization summary. Two persons worked on writing down the business idea description and the supporting text of BMC. Moreover, one of them worked on the hypothesis exposition and consequences while also gathering answers on the survey.

During our fourth meeting we met up to deliver, evaluate our work and make any changes required. After discussing, we made some changes and polished out our work overall. Finally, we put all the parts together into a single report.

References

- O'Driscoll, D. (2018). Time-series forecasting of academic careers. *Masters Dissertation*.