Optimal Stopping Algorithm for Modern Interview Process using Machine Learning

Husain A. Shaikh[1] Ankit A. Pandey[2] Asim M. Raj[3] Premraj V. Zambre[4]
[1,2,3,4]Student [5]Assistant Professor
[1,2,3,4,5]Department of Computer Engineering
[1,2,3,4,5]Keystone School of Engineering, Pune, India

Abstract— This Research Paper Aims to Provide the Balance of Time and Quality in modern interview process using Optimal Stopping Algorithms coupled with machine learning algorithm called as decision tree classifier. In Any optimal stopping problem, the crucial dilemma is not which option to pick, but how many options to even consider. We try to answer this question by combining optimal stopping algorithm and machine learning with human intuition.

Keywords: Optimal stopping algorithm, Decision Tree Classifier, 37% rule, Interview Process

I. INTRODUCTION

In Any optimal stopping problem, the crucial dilemma is not which option to pick, but how many options to even consider. These problems turn out to have implications not only for lovers and renters, but also for drivers, homeowners, burglars, and beyond.

The 37% Rule derives from optimal stopping's most famous puzzle, which has come to be known as the "secretary problem."

The establishment of the secretary problem is as follows. Suppose you are a Hiring manager who has a task to hire a secretary for your boss. You take up the challenge but face a problem later on with number of candidates. As candidates who are interested in the job role are quite large in number it might take quite a lot of time to interview everyone.

We might employ following two approaches:

1) Interview everyone and select the best among them or
2) Select the first candidate who meets the requirements

The problem with first approach is it takes quite a lot of time to interview everyone. While your company might stop at nothing to get only the best employees it might be challenging for small companies or startups to spend that amount of time on hiring.

The second approach then seems promising and intuitive. It makes sense to select the one candidate who meets the requirement at the earliest. But just as first approach it has a fundamental flaw. The flaw can be emphasized by following questions-

1) How do we rank them to appear for the interview?
2) How can we be sure that this candidate is best among all the candidates?
3) Can we take some more interviews to acquire a better suited candidate?

Above questions are harder to answer than we think. As it is apparent that our first approach which is to take every interview and then select the best performer takes a lot of time, while the second approach which is to select the first candidate who meets the requirements compromises with quality.

Even if we take n-1 interviews we cannot ignore the possibility of nth candidate being better suited than everyone else.

The balance of time and quality is an important problem that we face also in our daily lives. In the modern interview process it becomes necessary to handle this situation carefully as the candidate pool is quite large.

II. LITERATURE SURVEY

This section of the literature survey eventually will reveal some known facts based on the thoughtful analysis of the author's works as follows. While writing this research paper, there are more than 170 papers available on Google Scholar referring to the famous Optimal Stopping Algorithm or Secretary Problem & its extensions.

Brian Christian [1] in his book "Algorithms to Live By" says Optimal stopping is the science of serial monogamy. Optimal stopping algorithm offers solutions not only for the search for an apartment, but also for all those life situations in which we are faced with the question of an optimal stay. People face these problems every day, although the poets certainly spilled more ink on the court's difficulties than on the parking lot, and in some cases with considerable anguish. But anguish is not necessary. At least mathematically, these are solvable problems. The Optimal Stopping algorithm tells the balance is 37%.

According to the Researcher Zishuo Zhao [2], the "37% balance" for Optimal Stopping Algorithm has largely influenced people's opinion on decision making strategies about making a choice. In his research paper he mentioned as an online decision problem, Optimal Stopping Algorithm requires a Decision Maker to go through a series of n candidates and then decide whether to hire or immediately reject after observing one candidate, and the end goal is to maximize the chances of selecting the best candidate (Ferguson 1989 [3]). In the original problem, the ideal solution is to skip first n/e candidates and hire one immediately if it is best. The maximized probability is simply 1/e. Meanwhile author also suggests, in general, it is often more desirable to choose a "good" applicant, not precisely the "best".

According to Ferguson [3] optimal stopping algorithm or secretary problem has a certain appeal. It is easy to state and has a conspicuous solution. It was instantly taken up and advanced by certain well-known statisticians like Lindley, Dynkin, Chow, Moriguti, Robbins & Samuels, and Gilbert & Mosteller. From that time, the Optimal Stopping Algorithm or Secretary problem has been developed and became widespread in many different areas so that now one can say that it establishes a field of study within statistics-probability-optimization. One can read from the review paper by Freeman(1983) how extensive and vast

the field has become, furthermore, the field has sustained its exponential growth in the years since that paper appeared.

As stated by Theodore Hill [4] in his research paper, every decision humans make involves risk factors. Selecting the best time to stop and hire a candidate is critical. Bad timing can be critical. We face negligible consequences while stopping decisions all time, when responding to a job offer, hunting for better housing prices. The basic framework of all these Optimal Stopping problems is same. A decision maker observes a process that evolves over time and is somewhat random. Based on what is known, a person must take decision on the way of maximizing the reward or minimizing cost.

III. PROBLEM

How do we balance time and quality in the modern interview process?

As stated earlier, we will be employing optimal stopping algorithm more specifically the 37% rule along with Machine learning and common intuition to solve this problem.

IV. SOLUTION

Many of the thinkers earlier have tried to automate the process of hiring. Although it seems like we are progressing in terms of technology at a pace at which everything seems possible to automate the simple answer to automated hiring is "We are not there yet."

While taking an interview the interview panel looks for their requirements, candidate's experience, interests, energy, confidence, ability to express ideas clearly, etc. apart from their aptitude scores.

These analog entities cannot be translated to a single value so that a bot can take an interview and select candidates who meet the requirements.

Therefore, in our research we arrived at the conclusion that current interview technique is better - at least for now.

In this paper we will optimize this current interview technique which involves taking interview face to face by a panel and based on intuition select a candidate.

First, we will employ optimal stopping algorithms more specifically the 37% rule.

A. 37% Rule:

This approach caters to those who don't have much time despite that they would like to select the best candidate. 37% rule suggests that we look for at least 37% of the choices and then think about hiring. Although the original rule doesn't allow for going back and selecting a candidate who was rejected earlier, but we don't take this consideration into our implementation.

B. Machine Learning:

Although we have implemented 37% rule to optimize the process, we need a helping hand for the panel or a single interviewer that can suggest efficiently based on data that he has whether he should select or reject this candidate. We have experimented with different machine learning classifiers and found decision tree classifier to be the most efficient in terms of time and number of candidates. Although we are assuming candidate pool is large, it might not be always the case. Therefore, decision tree classifier caters our need when it comes to small training set. It's accuracy only improves as we add more candidates to the pool. It takes attributes and their values while the interview is in progress. As we've made interviewer to be the central entity of decision, the interviewer will rate each candidate in terms of their qualities based on organization's requirements, some qualities/skills might be Speaking fluency, clarity of thoughts, effective communication, leadership traits, etc. Based on this 37% of the candidates will be interviewed and interviewer will provide his/her decision to the algorithm as label for training. This training set will be used to then predict whether the candidate should be selected or even interviewed based on available data prior to and after the interview.

C. Human Intuition

As stated earlier human decision maker is the central decision maker, we will not give full command to the machine learning algorithm.

Although the classifier will provide suggestion in binary (0 or 1) whether we should select or reject the candidate, the final decision will be of the interviewer.

As we are dealing with someone's entire career, we should deal this situation with utter care as to take unbiased decision and record the parameters based on which the decision was taken.

V. IMPLEMENTATION

Implementation will be using python programming language on Django framework for GUI because after years of experience with different web technologies, we believe Django framework offers more features as compared to the other frameworks available out there. Machine Learning Algorithm will be provided by the scikit-learn Library which is commonly known as sklearn. A simple algorithm is used to implement 37% rule.

Storage of data will be done by using PostgreSQL for its simplicity and flexibility.

VI. CONCLUSION

The Optimal Stopping Algorithm has received large attention by statisticians and applied mathematicians all over the world. One reason for this is the problem's amazing optimal policy. Under it, the Decision Maker skips the first n/e candidates and then takes the next applicant whose value is a maximum. What is surprising is that in n/e, the $n \to \infty$, and that in the limit the best overall candidate is selected with probability 1/e.

We began this paper by presenting Optimal Stopping Algorithm for Modern Interview Process. The problem introduced in the current paper has similar intuitive solution. The balance of time and quality is an important problem that we face also in our daily lives. In the modern interview process it becomes necessary to handle this situation carefully as the candidate pool can be quite large.

Future research will focus on the actual implementation of Optimal Stopping Algorithm for Modern

Interview Process using Machine Learning in generalized Optimal Stopping problems. Compared to Classical Optimal Stopping Algorithm Problem, it seems to us that the solution mentioned above has more scope.

ACKNOWLEDGEMENT

It gives us great pleasure to present this paper on Optimal Stopping Algorithm for Modern Interview process Using Machine Learning. We would like to thank our research guide and Project coordinator Prof. Jubber S. Nadaf, Assistant Professor, Department of Computer Engineering, Keystone School of Engineering, Pune for their constant encouragement & valuable guidance during the completion of this research activity.

We are also grateful to Prof. Vilas G. Mankar, Head of Department, Department of Computer Engineering, Keystone School of Engineering, Pune for his constant support and valuable suggestions.

REFERENCES

[1] Brian Christian and Tom Griffiths, "Algorithms to Live By: The Computer Science of Human Decisions", ISBN 9781250118363
[2] Zishuo Zhao, "An Asymptotic Analysis on Generalized Secretary Problem", Computer Science and Game Theory (cs.GT), 16 Sep 2019
[3] Thomas S. Ferguson, "Who Solved the Secretary Problem?", Statistical Science, Vol. 4, No. 3 (Aug., 1989), pp. 282-289
[4] Hill, Theodore P. "Knowing When to Stop: How to Gamble If You Must—the Mathematics of Optimal Stopping." American Scientist, vol. 97, no. 2, 2009, pp. 126–133.

