Education & Career Success

Tran Dang Huy, Sia Chawla Chawla, Jiaqi Xu

Table of contents

1 Executive summary

(Maximum of 4 sentences)

2 Introduction

(Maximum 10 sentences)

Graduates today face many questions about how their experiences influence their careers. Grades alone may not determine who gets the most job opportunities. This report explores which student experiences are linked to receiving more job offers?

To answer these questions, we use a dataset of 5,000 recent graduates from Kaggle. It includes information about students' academic background, personal demographic, and career-related outcomes. Rather than testing predefined theories, this project takes an open-ended, pattern-oriented approach. The aim is to explore which types of experiences appear most consistently linked to job outcomes and personal satisfaction.

Table 1: Dataset variables and their descriptions

```
Rows: 4,808
Columns: 20
$ Student_ID            <fct> S00001, S00003, S00004, S00005, S00006, S00007, ~
$ Age                   <int> 24, 28, 25, 22, 24, 27, 20, 24, 28, 28, 25, 22, ~
$ Gender                <fct> Male, Female, Male, Male, Male, Male, Male, Male~
$ High_School_GPA       <dbl> 3.58, 3.42, 2.43, 2.08, 2.40, 2.36, 2.68, 2.84, ~
$ SAT_Score             <int> 1052, 1193, 1497, 1012, 1600, 1011, 1074, 1201, ~
$ University_Ranking     <int> 291, 715, 170, 599, 631, 610, 240, 337, 138, 594~
$ University_GPA        <dbl> 3.96, 2.63, 2.81, 2.48, 3.78, 3.83, 2.84, 3.31, ~
$ Field_of_Study        <fct> Arts, Medicine, Computer Science, Engineering, L~
$ Internships_Completed <int> 3, 4, 3, 4, 2, 0, 1, 2, 1, 2, 2, 2, 0, 1, 3, 3, ~
$ Projects_Completed    <int> 7, 8, 9, 6, 3, 1, 5, 3, 5, 7, 2, 0, 4, 2, 5, 2, ~
$ Certifications        <int> 2, 1, 1, 4, 2, 3, 5, 0, 3, 5, 3, 5, 3, 3, 2, 3, ~
$ Soft_Skills_Score     <ord> 9, 1, 10, 10, 2, 3, 5, 5, 10, 8, 2, 2, 8, 1, 9, ~
$ Networking_Score      <ord> 8, 9, 6, 9, 2, 3, 1, 5, 2, 1, 9, 9, 6, 8, 9, 5, ~
$ Job_Offers            <int> 5, 0, 1, 4, 1, 2, 2, 2, 0, 5, 5, 2, 2, 1, 3, 3, ~
$ Starting_Salary       <dbl> 27200, 42400, 57400, 47600, 68400, 55500, 38000,~
$ Career_Satisfaction   <ord> 4, 9, 7, 9, 9, 7, 2, 2, 4, 9, 7, 9, 4, 9, 7, 8, ~
$ Years_to_Promotion    <int> 5, 3, 5, 5, 2, 4, 3, 2, 2, 1, 4, 4, 3, 3, 4, 5, ~
$ Current_Job_Level     <ord> Entry, Entry, Mid, Entry, Entry, Mid, Entry, Ent~
$ Work_Life_Balance     <ord> 7, 7, 5, 2, 8, 3, 3, 2, 2, 2, 6, 8, 3, 6, 3, 9, ~
$ Entrepreneurship      <fct> No, No, No, No, Yes, No, No, No, No, No, Yes, No~
```

3 Methodology

(Maximum 300 words; Should include a figure and a table and those must be referenced in the text and have adequate captions)

This study uses an exploratory approach to investigate how students' academic and experiential attributes may relate to early career outcomes. The dataset includes 5,000 graduate records collected from Kaggle.

We began by examining the structure of the input variables. As shown in Table 1, the dataset contains a mix of continuous (e.g., `SAT_Score`), ordinal (e.g., `Certifications`), and categorical (e.g., `Field_of_Study`) features, with a wide range of unique values. This diversity limits the applicability of simple linear models and supports the use of flexible methods that can accommodate mixed data types and non-linear relationships.

We also analysed the outcome variable. As visualised in Figure 1, variable `Job_Offers` is relatively well-distributed and show minimal skewness.

Given that `Job_Offers` is a count variable, we select random forest model, a non-parametric model capable of capturing complex relationships without strong distributional assumptions.
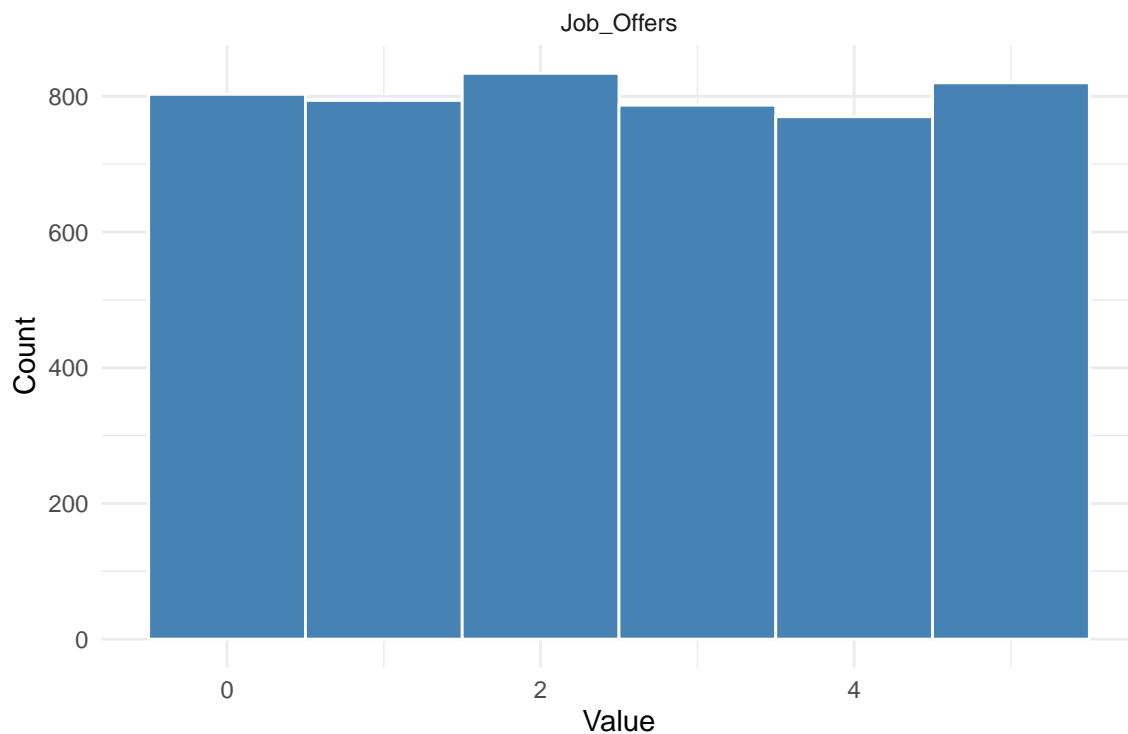
Figure 1: Distributions of Outcome Variable

4 Results

(Maximum 200 words. Should include either a figure or a table.)

4.1 Associated factors for receiving job offers.

Figure 2 shows the relative importance of predictors in the random forest model. Variables such as number of complete internships and projects, SAT score, and gender were most associated with the number of job offers receiving by the student. Negative importance scores (e.g., for Age) suggest minimal or noisy contributions to model performance.

4.2 Association between important factors and job offers.

5 Discussion, conclusion and recommendations

6 Reference section
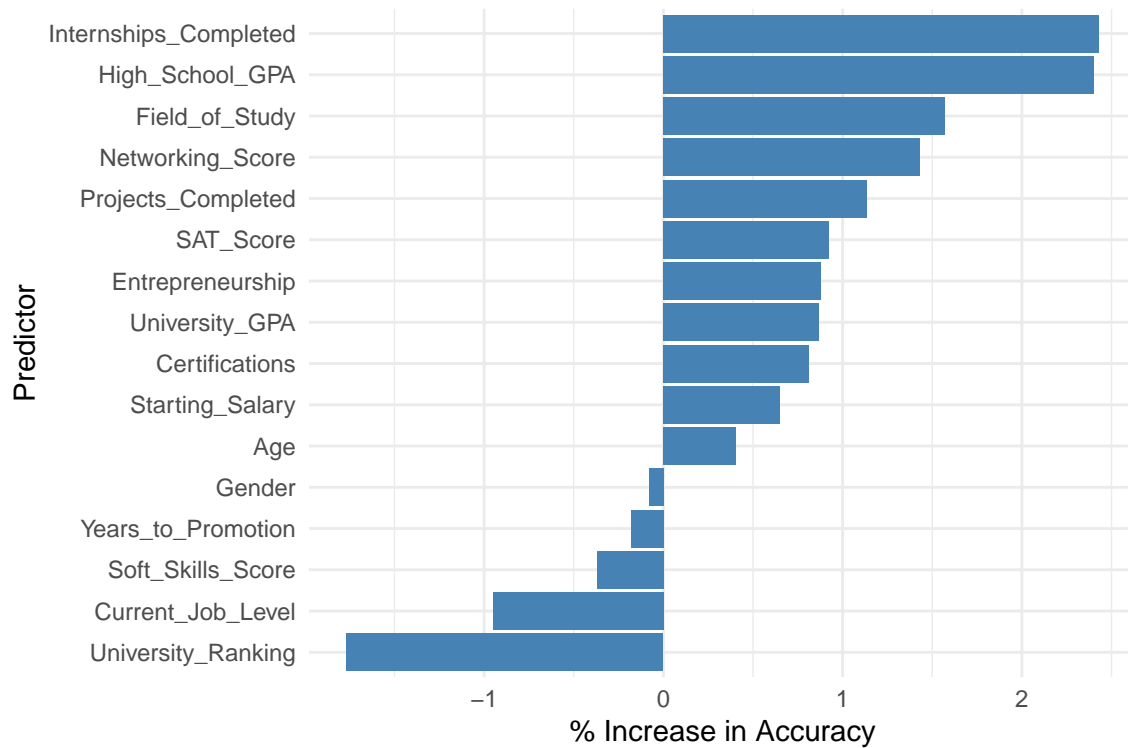
(Include at least 1 reference.)

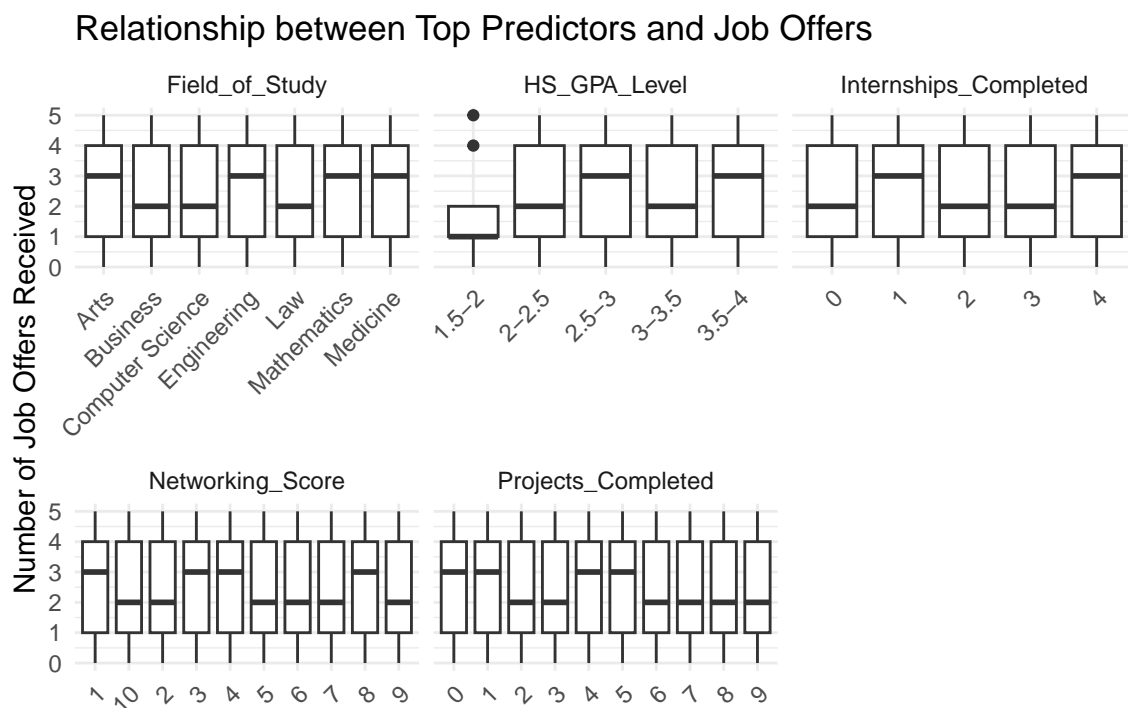Figure 2: Importance of Predictors for Number of Job Offers Received

Figure 3: Relationship between the most important student experiences and job offers received