**Using Statistical Learning and Big Data to Predict Relationship Status**

*A Case Study of Middlebury College Students*

Quinlan McGaugh ’22 and Otis Milliken ‘24

Middlebury College

**Abstract**

**Introduction**

* Intro and Motivation

If you have taken a tour at Middlebury College, odds are high that you heard the tour guide boast an interesting statistic regarding marriage rates of Middlebury Students. Something like, “You know, something like 60 percent of Middlebury students marry other Middlebury students![[1]](#endnote-1)” Although some questions have been raised, and these relationships are a key piece of college life. During Covid-19, relationships were

Motivated by this interesting phenomenon, as well as series of polls like the Middlebury Campus’ *Zeitgeist 4.0* and Midd MarriagePact, the authors aimed to create a tool for predicting

* Why is the problem important
* Novel contribution to the field of AI

Section II will discuss a brief literature review of random forests and decision trees. Section III will formalize our problem statement. Section IV will provide a full description of our methodology and results. Section V offers conclusions and discussion.

**Related Work**

**Problem Statement**

**Methods and Results**

*I. Data*

*II. Statistical-Based Learning using Decision Trees and Random Forests*

*II.A Naïve Decision Tree*

*II.B Random Forests*

*III. Results*

*III.A Naïve Decision Tree*

*III.B Random Forest A*

*III.C Random Forest B*

*III.D Comparison of Techniques*

**Discussion and Conclusion**

**Bibliography**

**End Notes**

1. <https://middleburymagazine.com/features/bye-bye-love/> [↑](#endnote-ref-1)