**Making Up for Missed Connections in the Pandemic:**

**Machine Learning for Online Dating Scams**

1. **Abstract:**

Since the start of the COVID-19 pandemic, the use of dating applications and dating sites has increased greatly. According to the Wall Street Journal, eight of the largest dating apps used in the US had a 12.6%-year increase in monthly active users in 2020. Online dating site OkCupid saw a 700% increase in dates scheduled in 2020. As the number of users of online dating applications and dating sites increases, so do online dating scams. The Federal Bureau of Investigation reported a $1 billion loss from romance scammers in 2022 in the US. The ability to detect these scammers before they steal from their victims is critical to preserving the business of the dating application or dating site and the mental wellbeing of the victim. Companies of dating applications and dating sites currently have scam investigation teams and preliminary detection algorithms, but the nature of online dating scams make it particularly challenging to detect. The goal of this research is to create a scam probability index for online dating profiles through a machine learning algorithm to enhance scam detection methods. The aim is to create this index as an aid to scam investigation teams and flag potential scam profiles early on to prevent any loss.

1. **Introduction:**

Recent studies show that scammer dating profiles tend to act differently than the average real-life user. Dating scams are usually run on a small scale and they develop overtime to gain trust with their victims. An online dating scammer is someone that uses the service to take advantage (most frequently economically) of another user. Common characteristics of a scam account include a large number of conversations initiated, little time between profile creation and the first message sent, and the large portion of messages received that the account replies to. Some ways to identify scammers are through an IP-Address-based Detection System, a Photograph-based Detection System, and a Text-based Detection System.

Beyond these attributes, specific types of online dating scams have specific goals and characteristics. Common types of dating scams include Escort Service Advertisements, Swindlers, and Dates for Profit. Escort Service Advertisements are usually young attractive females that will forward information to their escort service agency when they reach out to other profiles. Swindlers seek to establish trust with a real user, ask for money, then proceed to never message the victim after the money has been secured. Dates for Profit are commonly young females who go on expensive dates to earn money for the date location; they receive compensation, then never go on a second date with the victim.

This project aims to use machine learning to analyze all of these factors to create a relative scam index. The algorithm will assess profiles for the common scam features, as noted prior, and for the different types of scams.

1. **Methods:**

The datasets we will be using in our analysis are from available open source websites that contain real profile data and scam profile data. The data ranges from different online dating sites to provide a holistic approach to scam profile characteristics. By comparing the scam profiles and the real profiles, we will add on to the common features of scam profiles as noted by previous studies. These common features will be included in our index assessment within the algorithm.

1. **Timeline:**

Week 1:

Week 2:

Week 3:

Week 4:

Week 5:

Week 6:

Week 7:

Week 8:

Week 9:

Week 10:

1. **Feasibility:**

There has been previous research on the variety and nature of online dating scams. Probability scales of scam dating profiles have been created prior, but they take a different approach and neglect the attributes of specific scams. The ambition of this project is to use machine learning to discern the type of dating scam along and assess the common characteristics of scam profiles.

1. **Broader Impacts:**

With the number of dating profiles increasing, it is imperative that online dating scam profiles be minimized to provide dating application and dating site users with a positive experience as they strive to make romantic connections. The COVID-19 pandemic negatively affected individuals’ ability to find prospective partners and spark their romantic life. Through this research, we would be able to help close the gap any romantic connections lost. Dating applications and dating sites would financially benefit from this research by cutting down on lost users who experienced a scam.

1. **Trajectory and Future Goals:**

As a History minor, it is important to me to use my technical skills as a means of positive social impact. Within the Computer Science field, I am interested in data analysis and using programming in research. I plan on pursuing graduate education whether that be for a masters or a doctorate degree. This project would help provide me with hands on research similar to that in a graduate setting. This hands-on experience would help distinguish me as a candidate for any graduate program. I am not confident on exactly what field of computer science I would like to pursue; thus, this project will help me determine whether data analysis and research is truly right for me.