MATH5925 Project Proposal

Predicting attempted security breaches at the Ministry of Magic

by Stats-Wiz:

Harry Potter (z123456)
Hermione Granger (zID)
Ronald Weasley (zID)

Submitted in partial fulfilment of the requirements of the project course MATH5925

Predicting attempted security breaches at the Ministry of Magic

Stats-Wiz
Our mission is to develop and apply state-of-the art statistical techniques to key problems to make the wizarding world a safer place for all witches and wizards.

Background

After the Tower of London, the Ministry of Magic has the highest security of any location in the United Kingdom. However, it is also one of the most challenging of locations to keep secure. On any give day, over ten thousand visiting witches and wizards can be expected to arrive by floo powder, broom, or any number of secret entrances in central London [1]. This has led to a number of ministry breaches, some of the more well-known being when Grindlewald was at the height of his powers [2], and a full scale battle between followers of He Who Must Not Be Named and the Order of the Phoenix [1].

In this project we aim to predict the likelihood of security breaches at the Ministry of Magic in real time. Security forces, while substantial, are a finite resource, and a predictive model could be used to inform when and how to optimally allocate these resources to minimise the risk of security breaches. Given that the Ministry houses key decision-makers in the wizarding world, and many highly sensitive documents, these efforts are much-needed and could have immeasurable benefit.

Regrettably, there has been no previous research on this topic. This is in part because we are talking about a fictional location, and there is only so much you can find out from the Harry Potter Wiki (`https://harrypotter.fandom.com/`). However, extensive literature on techniques for modelling cybersecurity breaches exist in the Muggle world [3, for example], and we intend to lean on that literature in developing our approach.

Objective

Our objectives are:

1. To construct a real-time predictive model for chance of a security breach at the Ministry of Magic.
2. To test its performance by forecasting attempted breaches in 2021, and through field testing in the Misuse of Muggle Artefacts Office.

Scope

This project will produce a website that reports in real time the likelihood of a security breach at the Ministry of Magic, using a green-yellow-red alert system, and a 200 page report on its underlying statistical model, intended usage, and recommended counter-spells. We anticipate that the website will be active within six months of commencement of this project. We will present the system to key Ministry personnel at the time. The system will then undergo operational testing for a further six months at the Misuse of Muggle Artefacts Office.

The model will be constructed based on Ministry provided data (from 1634 onwards) on:

- Time, location, nature and outcome of attempted security breaches (successful and unsuccessful)
- Number of magical persons of interest at large and their last known location
- Number and location of unauthorised uses of magic and magical items in the United Kingdom.

and the following data from the Muggle world:

- Date and time
- Population size
- Weather forecast and its congruence with actual weather
- Time series for the Financial Times Stock Exchange 100 Index since 1984

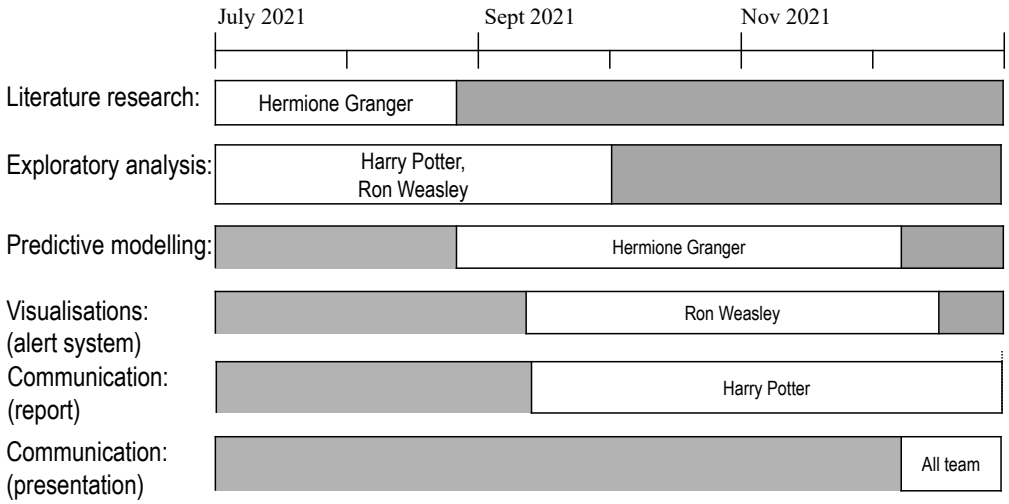
In order for this model to have validity for forecasting purposes, we will need to assume that the factors affecting likelihood of security breaches at the Ministry of Magic do not change in the future from what they have been historically. We further assume a stationary model, in the sense that there has been no change at all historically in the effect of these factors on security breach attempts. While the Ministry has always taken great effort to ensure standardised data collection practices, there have been some changes in practices over the last four centuries, and we will need to assume that these do not affect the nature of the association between predictors and our response.

In measuring the success of our predictive model, we expect that it will have high correlation with actual attempted security breaches in 2021 (an AUC of at least 0.9). We will then trial our system in the Misuse of Muggle Artefacts Office, where we anticipate that using a flexible security staffing pattern in response to risk will reduce the chance of successful breaches by 10%.

The ongoing COVID19 situation has reduced our access to key Ministry personnel and locations. We now have access to all relevant data and can proceed with model-building, but a further outbreak could delay the operationalisation of our model at the Misuse of Muggle Artefacts Office. This matter is regrettably out of our control but we will discuss COVID19 projections with Ministry staff on a monthly basis and automatically delay operationalisation by three months in the event of a lockdown in late 2021.

Project Plan and Timetable

Below is a timetable for Phase I, building the predictive model, and forecasting 2021 breaches. Field testing will be considered at a later stage.



The key steps in Phase I of this project are as follows:

Literature research (6 weeks), investigating methods of predictive modelling suitable for this dataset. Hermione Granger is an experienced scholar who is ideally positioned to lead this work.

Exploratory data analysis (8 weeks), which will involve cleaning data, resolving errors in the data in consultation with the Ministry, and constructing plots to explore core properties. Early results suggest that a key predictor of incidents is time of day, with attempted breaches most often occurring in the early morning, especially on Saturdays after major Quidditch matches. Ron Weasley and Harry Potter will lead this work, Ron having previous experience with Ministry data and working with Ministry personnel, and Harry having extensive experience with the Muggle world and its bureaucracy.

Predictive modelling (12 weeks), led by Hermione Granger (Harry Potter to assist). She plans to build a convolutional neural network to predict likelihood of a security breach, and its outcome, using the VGG-19 architecture [4], as a function of all available predictors. She will build the model using data up to 2020, optimised on one-day forecast success, and will later use it to construct one-day forecasts for 2021 of breaches and their outcomes. She has not applied convolutional neural networks before but has extensive programming and analysis experience.

Visualisations (10 weeks), with a special focus on developing an intuitive alert system for Ministry personnel to use. Ron Weasley will build such an alert system to minimise mean misclassification error on one-day forecasts. Ron's previous experience working with Ministry personnel will assist in making this system accessible to Ministry employees.

Communication (12 weeks) includes: a report (expected length: 200 pages) led by Harry Potter; and a group presentation (from the whole Stats-Wiz team). All team members will contribute to report writing, Harry will co-ordinate.

References

[1] C. Binns, Modern Magical History, Rumihart London, 2020.

[2] B. Bagshot, A History of Magic, Little Red Books, 1947.

[3] B. Edwards, S. Hofmeyr, S. Forrest, Hype and heavy tails: A closer look at data breaches, Journal of Cybersecurity 2 (1) (2016) 3–14.

[4] Y. Liu, W.-C. Lee, G. Tao, S. Ma, Y. Aafer, X. Zhang, Abs: Scanning neural networks for back-doors by artificial brain stimulation, in: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, 2019, pp. 1265–1282.