**RESEARCH PROPOSAL – Predictive Analytics and Study Abroad**

Blake Ward

9/13/17

[blake.n.ward@gmail.com](mailto:blake.n.ward@gmail.com) (414-336-67930

**Purpose:**

One of the primary goals of the Office of International Education at Marquette is to increase the rate of study abroad participation across the undergraduate population. Although the office has successfully maintained a consistently high study abroad rate relative to peer institutions, a significant percentage of the over 8,000 undergraduate students still leave Marquette without having a meaningful, credit bearing international experience. This project aims to assist OIE by allowing the staff to potentially target students on the margin of studying abroad, those who might be convinced to take part if they were properly informed of the options and benefits. It also aims to help OIE in its advising by providing advisors a better understanding of the type of students who choose to study abroad and those that do not, along with an idea of what type of students choose certain locations and how they can be advised on the best options.

**Methods:**

Although specific algorithms will be chosen after exploratory analysis of the available data, the general method will be to choose a supervised learning algorithm (logistic regression, Naïve Bayes, random forest, etc.) that probabilistically classifies students as those who are likely to study abroad and those who are not. A similar technique may also be used on location preference for those that choose to study abroad. Also, an unsupervised clustering algorithm will be used to identify patterns in the cohort of students that study abroad to allow for more targeted advising and marketing.

**Requested Data:**

Listed below is a complete list of data elements requested for the study (obviously dependent on availability). Ideally, the datasets would be provided as a snapshot as of the end of each year (spring 2013 to spring 2017) on all enrolled students. If point-in-time snapshots are not available, the GPA for each student as of the end of each school year is requested along with the most current data for the other elements. If that is not available, the most current data for each element in the “snapshot” section is requested. The time range requested is the population of students enrolled from Fall 2013 until Spring 2017, including the graduating class of 2016-17.

To meet FERPA standards, data that is de-identified is requested. In order to link datasets from OIE and OIRA, a standard process of scrambling the student IDs would be appreciated (hashing, etc.). No other personally identifying information is requested. In the case where only a student email can be used to link datasets, any standard method of scrambling the email so that they can be joined would be ideal. If that is not possible, the survey data may not be able to be used.

**OIRA Elements/PeopleSoft Data**

Demographic/Background Data

1. Student ID (scrambled)
2. Student email (scrambled – only if necessary to link to survey data below)
3. Gender
4. Ethnicity
5. Birthdate
6. City/State of residence (i.e. hometown)
7. High School
8. High School GPA
9. Class Rank
10. Entrance Exam Results (ACT and/or SAT)
11. Pell grant eligibility at time of application
12. Amount of MU financial aid offered upon admission (per year)
13. Academic Advisor
14. QPA (predicted GPA)

Snapshot/Point-in-Time Data (ideally end of each academic year)

1. Student ID (scrambled)
2. MU scholarship amount received (for the year)
3. Academic major/minor
4. Academic standing (freshman, sophomore, etc.)
5. GPA in major
6. GPA in minor
7. Taken a service learning course (yes/no)
8. Participated in a student club/extracurricular (yes/no)

Survey of First Time Student Data

1. Student ID (scrambled – or email)
2. Indicated interest in cross-cultural experience (yes/no)
3. Indicated interest in international service learning (yes/no)
4. Indicated interested in study abroad (yes/no)

OIE Data

1. Student ID (scrambled)
2. Study Abroad Year
3. Study Abroad Semester
4. Study Abroad Duration (Semester, Summer, etc.)
5. Study Abroad Location (City, Country)
6. Study Abroad Type (Short-term MU, 3rd Party Provider, Partner Program, etc.)

**Analyst Bio**

Blake Ward is a former International Communications Coordinator and Study Abroad Coordinator at Marquette. Currently working as a Big Data Engineer at Johnson Controls, he has also held analyst and data engineering positions for the Milwaukee Public Schools. This proposed work is part of a capstone project for a Machine Learning nanodegree program through Udacity, and also to help former colleagues in an area that still holds a keen interest for Blake.