

Progress Report

By Alex Hoffer, Jake Smith, and Chen Chongxian

Team Name: Stat Champs

Abstract

The application of machine learning to Biochemistry and Biophysics has enabled researchers in this field to make remarkable discoveries, such as the generation of new DNA sequences. However, students of Biochemistry and Biophysics do not get the opportunity to learn machine learning. Dr. Victor Hsu of the Oregon State University Biochemistry and Biophysics department has commissioned the Stat Champs to produce an instructional module to give his students the chance to familiarize themselves with machine learning. The software product the Stats Champs have agreed to develop is a web page that allows students to train a machine learning model based on the college basketball statistics and machine learning algorithm of their choosing in order to produce a March Madness bracket. This will help students understand how machine learning algorithms produce models and how inclusion or exclusion of certain data can influence such models. Over the course of Fall term 2016, the Stat Champs have developed materials such as design documents and technology reviews in order to prepare for the engineering of this module. This report comprehensively describes the progress the Stat Champs have made thus far.

Table of Contents

1	Introduction	3
2	Progress Descriptions	3
2.1	Alex Hoffer	3
2.2	Chongxian Chen	4
2.3	Jake Smith	4

1 Introduction

This report chronicles the progress the Stat Champs have made on developing the machine learning instructional tool. As of the end of Fall term 2016, the progress that has been made consists of a project statement, a requirements document, a design document, a technology review, and this progress report. The remaining sections of this report are devoted to each of the individual members of the Stat Champs to describe their experiences this term as well as their contributions to these documents.

2 Progress Descriptions

2.1 Alex Hoffer

In week one, we were assigned to the "Machine Learn Your Way to March Madness Glory" project proposed by our client Dr. Victor Hsu. During week two, we met with Dr. Hsu to get clarity on the specific details of the project. A synopsis of the project can be found in the Abstract of this document.

After meditating on what we learned from our exchange with Dr. Hsu, we wrote a rough draft of our project statement. In week three, we refined this document and sent it to Dr. Hsu for his approval. The project statement consisted of an abstract, a problem definition, a proposed solution, and performance metrics. We were all relatively new to LaTeX and so we had some issues formatting this document correctly. Dr. Hsu felt that we needed to change several things in the document, so in week four, we made these changes and re-submitted it.

During week five, we began developing our project requirements document. I made individual progress by writing the document in Word. This document was formatted using the IEEE standard and thus consisted of an introduction, an overall description, specific requirements, and a Gantt chart. My individual contribution included writing the introduction, an overall description, and the specific requirements. The next week, we adapted my writing to a LaTeX file developed by Jake and added a Gantt chart that Chongxian made. Then, we submitted the final product to Dr. Hsu and got his approval.

When week seven came, we revised and re-submitted our project requirements document because Dr. Winters felt it was too vague. Individually, we made progress on our sections of the technology review document. In week eight, we each individually completed our technology review sections and compiled them into one large LaTeX formatted document. We faced some difficulty in coming up with responsibilities for each of us and realizing what technologies would be best suited to satisfy said responsibilities. My responsibilities were centered around the usability and appearance of the system. They were 1) the graphical user interface of the web page 2) a presentation of instructions to the user of the module and 3) the presentation of the machine learned bracket. After careful consideration, I landed on using PyGUI for 1), Webix for 2), and VTK for 3).

For week nine, we discussed as a team what approach we should take to the design document, and each individually landed on different software design templates. We also discussed how we would record ourselves for the progress report. One solution we considered was renting out a microphone from the library. As usual, we had some issues with LaTeX formatting.

In the final two weeks, we finished and submitted our design document. I used sequence diagrams to describe the functionality of the technologies I selected. In my rationale, I included the average use case for each of my responsibilities. It was difficult getting LaTeX to properly generate the sequence diagrams, particularly because the packages I installed were being stubborn. Then, we got together and recorded a progress report. Finally, we each wrote our section of this document and compiled it together.

Retrospective:

Positives	Deltas	Actions
Finished project statement and requirements document	Need to learn basic Python	Will take Python tutorials over winter break
Finished design document and progress report	Need to learn basic JavaScript	Will take JavaScript tutorials over winter break
We all learned LaTeX	Need to learn machine learning algorithms	Will check online to understand the machine learning algorithms Chongxian wants to use
Prepared for engineering phase	Need to learn my selected technologies	Will read the documentation for my technologies over winter break

2.2 Chongxian Chen

2.3 Jake Smith