



**CALIFORNIA STATE SCIENCE FAIR
2017 PROJECT SUMMARY**

Name(s) Christopher L. Tan	Project Number S0834
Project Title Using Machine Learning to Predict the Flu	
Abstract Objectives/Goals Use machine learning to prevent illness and prevent epidemics by predicting strains of H1N1 and H3N2 influenza viruses. Methods/Materials Accessed online databases to get large quantities of amino acid sequences for hemagglutinin and neuraminidase proteins, researched characteristics of the influenza virus as well as the surface glycoproteins, and used that information to create predictive algorithms in Python to produce possible epidemic-inducing strains of the influenza virus. Results Created machine learning algorithms to predict possibly infectious strains of influenza virus. There was little difference between the additive and the multiplicative algorithms, which both conserved parts of amino acid sequences while allowing enough variability for the antigens (the surface glycoproteins) to be effective. By testing similarity to common strains from previous years, I determined the efficacy of each algorithm. Conclusions/Discussion Although I was not able to test the strains my machine predicted in reality, the results suggested that the strains my computer produced using the machine learning algorithms could be viable and possibly predictive, since they were able to conserve parts of the amino acid sequences while changing enough to alter the antigens, preventing the immune system from recognizing them and thus allowing them to infect us. This project demonstrates the possible applications of bioinformatics to prevent illness, and specifically, how machine learning can be applied to diseases that change rapidly, such as viruses, to help predict their mutations.	
Summary Statement I created machine learning algorithms to predict amino acid sequences of surface proteins for H1N1 and H3N2 influenza viruses.	
Help Received Mr. O'Shea and Mr. Heyward suggested viable ways to expand original project. I conducted research and created machine learning algorithms on my own.	