California State University San Bernardino

School of Computer Science and Engineering

### CSE 695 Graduate Independent Study

**Date**

December 5, 2019

**Time**

1:00PM

**Location**

JBH 359

**Title**

Bioinformatics: Sequences and Genome Analysis

**Student**

Andrew Tran

**Advisor**

Dr. Ernesto Gomez

**Abstract**

Bioinformatics is an interdisciplinary field that combines concepts from biology and computer science in order to develop methods and tools in order to understand biological data. With the continual improvements to technology, the ability to understand and compile biological data becomes more increasingly available. In regards to the more technical aspects of Bioinformatics, the project will focus on the introduction of pair-wise and multiple sequence alignment of strings in relation to biology. This will be done through the introduction of various methods including the use of dot-matrix analysis and dynamic programming in order to compare sequences. These methods allow researchers to find matches and gaps between sequences, which then allow for prediction and trace-back of those sequences. In regards to the biological aspect, the project also covers the topics of collection and storage of sequences in laboratories and databases along with phylogenetic tree predictions in DNA, RNA and protein. Bioinformatics may be a relatively new field, but its applications, such as the (near)-completion of the Human Genome Project, continue to grow with each passing day.