Michael Lee

CSE 446

Project Proposal

Project title- “Predicting Gene Expression Levels across Cancer Afflicted Patients”

Data set- The Cancer Genome Atlas (TCGA)

Project ideas-

Predicting expression level of regulation genes across cancer afflicted patients

\*Individuals with a certain cancer? Across All cancers?

Individ 1: OncoGene1 | Tata Box Mutations | Other Mutations.. | Start Site Mutation|

Features:

TATA box mutations

Transcription Start Site Mutation

Histone Modifications

Software you will need to write.

ML Algorithm in google BigQuerry

**Papers to read.** Include 1-3 relevant papers. If you are doing something different then one of the suggested projects, you will probably want to read at least one of them before submitting your proposal.

<http://genomebiology.biomedcentral.com/articles/10.1186/gb-2012-13-9-r53>

<https://cancergenome.nih.gov/publications>

**Teammate**: will you have a teammate? If so, whom? Maximum team size is two students. One proposal per team.

I will be completing the project individually.

**Milestone**: What will you complete by the milestone? Experimental results of some kind are expected here.