**Title Page**

* Denny Shin
* November 2019
* Submitted in total fulfilment of the requirements of the degree of Master of Science (Bioinformatics)
* Department of Science (Bioinformatics)
* The University of Melbourne

**Table of contents**

**Abstract**

Do at end

**Declaration**

This is to certify that:

1. the thesis comprises only my original work towards the masters except where indicated in the Preface,
2. due acknowledgement has been made in the text to all other material used,
3. the thesis is less than 15,000 words in length, exclusive of tables, maps, bibliographies and appendices as approved by the Research Higher Degrees Committee.

**Preface**

* Where did data come from
* Some previous works by Alex

**Introduction**

Background

* chromatin accessibility
* machine learning chromatin accessibility

Significance

* eQTL
* non-coding variant analysis
* combinatory effect of motifs and its variants

**Objectives + Research Question**

Build and verify a neural network model that incorporates sequence and footprint information to predict the effect of non-coding genetic variants on chromatin accessibility

**Literature Review**

* CENTIPEDE
* centiSNP
* DeepBind
* deltaSVM
* Basset??

**Methodology**

* Creating the instances from CENTIPEDE data
* Training the models
* Evaluating and picking a model
* Creating new instances from centiSNP data
* Prediction of dsQTLs
* Verifying against dsQTL results

**Results**

* Model results
* Generated instances from centiSNP data
* Verification against dsQTL

**Discussion**

**Conclusion**