# SIT720 Machine Learning

## Task 8.1

Michael Rideout
Student Id: 225065259

# 1.1 Main Points Summary

## Week 7 Summary

Week 7 focused on the topic of SVM (Support Vector Machines). They can represent nonlinear functions and have been shown to achieve high accuracy in real life scenarios. SVM attempts to find the optimal hyperplane that maximises the margin between two classes.
When data is not perfectly separable, a soft margin can be introduced where some data points can fall outside the margin within a certain tolerance. The regularisation parameter C is the trade-off between maximising the margin and minimising misclassification errors.

SVMs can handle non-linearly separable data through a method called the ‘kernel trick’. The trick is to map the data into a higher dimensional feature space where it does become linearly separable. Similarity measures between points then occur in the higher dimension transformed space. Common kernel functions are linear kernels (create linear boundaries in the original space), polynomial kernel (creates curved boundaries) and Radial Basis Function (RBF, complex non-linear boundaries by mapping to an infinite-dimensional space)

SVM has a foundation in statistical learning theory. Structural risk minimisation is a method to balance the empirical error with the model complexity to prevent overfitting. Vapnik-Chervonenkis (VC) dimension measures the complexity of hyperplanes. Maximising the margin helps minimise VC hence leading to better generalisation.

Multiclass classification is achievable with SVM by using the following strategies:

- **One vs Rest** - Trains N binary classifiers for N target classes where each classifier distinguishes one class from all the rest. The class with the highest confidence score across all N classifiers is chosen
- **One vs One** - Trains a classifier for every combination of pairs of classes in the target. Final prediction is done by a voting scheme.

## Week 8 Summary

The topics covered in week 8 were KNN (K-nearest Neighbours) and decision trees.

KNN is a supervised learning method capable of both classification and regression. It represents instances in a feature space, and calculates the distance from that instance to all other instances. For classification, it finds the K nearest neighbours and then takes the majority of the target class of those neighbours. For regression, it takes the mean of the target of the neighbours. Some notable variations of KNN involve giving more influence to neighbours that are closer (Shephard's method)

There is a lot of optimisation that goes into finding the appropriate K value for KNN. Small K can lead to low bias and high variance while large K causes the inverse. Cross validation is often used to find the optimal K through the use of evaluation maximisation. Differing distance metrics can also affect the performance of KNN.

Decision trees were also addressed this week. They are a supervised learning model that can also be used for classification and regression. The idea is to construct a tree of nodes that represent a decision on variables in the feature set. The general steps in generating a tree are to start at a root node, find the best split for the node (divides dataset into two or more child nodes), iterate on every node until no further gain.

Regression decision trees typically determine how to split nodes on the basis of minimising the mean squared error.  Classification decision trees can use various methods to split nodes, however one of the most common is Gini Impurity which calculates how likely some random picked instances would be classified incorrectly.

Decision trees can be afflicted by model complexity if the depth or breadth of the tree grows too large. Tree pruning strategies exist to reduce the number of nodes in a tree that is deemed to be too large.

Some key advantages of decision trees are their interpretability and ability to model nonlinear functions. Some disadvantages are that they are prone to overfitting.

# 1.2 Summary of Reading List Items

## Week 7 Reading Items

**Video – Support Vector Machine - How Support Vector Machine Works**

This video explains SVMs. It explains the concepts of hyperplanes and support vectors. It describes what happens when data is not easily separable and introduces kernel functions.

**Video - The C Parameter for Support Vector Machines**

This video covers the C SVM hyperparameters and describes how it is a balance between a wide separation margin and maximising classification accuracy. High C values prioritise classification accuracy but with a narrower margin. Low C values are the reverse of this.

## Week 8 Reading Items

**Video - United States of Voronoi**

Uses a map of the United States to demonstrate Voronoi diagrams, a method of dividing a plane into regions or cells based on the distance to a set of points. It shows examples of classical problems that can be solved using Voronoi diagrams. Delani triangulation is a graph derived from a Voronoi diagram, which was also demonstrated.

**Video - Lecture 21: Regression Trees**

Covers regression trees, a way to model Y an outcome as a function of input variables X. The object of RT is to define regions that minimise the squared error within each region. The splitting algorithm is described and how splitting ends. Overfitting it also touched on.

## 1.3 Learning Reflection

From week 7 and week 8 lectures I have gained knowledge on the following machine learning algorithms, SVM, KNN and decision trees. For SVM, we learnt how data is separated between two classes and the optimal hyperplane between them is found. For KNN, we learnt how instances are represented in n-dimensional space and then using a distance measure, the closest instances are used for either classification or regression purposes. We learnt what decision trees are, how the are constructed (notable how nodes are iteratively split) and what mechanism exist to prune overly complex trees to aid in the reduction of overfitting.

# 1.4 Quiz Results

Week 7 Quiz:

**Week 7 Quiz**

Your work has been saved and submitted

Written 17 May, 2025 10:36 PM - 17 May, 2025 10:38 PM • Attempt 2 of unlimited

| | |
|---|---|
| Attempt Score | 9 / 10 - 90 % |
| Overall Grade (Highest Attempt) | 9 / 10 - 90 % |

Week 8 Quiz:

## Week 8 Quiz

# Your work has been saved and submitted

Written 17 May, 2025 10:45 PM - 17 May, 2025 10:51 PM • Attempt 1 of unlimited

| | |
|---|---|
| Attempt Score | **9 / 10 - 90 %** |
| Overall Grade (Highest Attempt) | **9 / 10 - 90 %** |