School of Computing and Information Systems
The University of Melbourne
COMP30027 Machine Learning (Semester 1, 2021)

Tutorial: Week 10

1. Hidden Markov Models (HMMs) are best used when the observables are a **univariate time series**: we are just observing a single variable, which changes over time due to some factor that can be estimated from previous observations.

   (a) Recall the two main assumptions (Markov, output independence) that are built into an HMM.

   (b) Could we construct the HMM in such a way to relax these assumptions? What would the model look like, and what is the major downside?

   (c) Could we build an HMM for a **multivariate time series**, where we have a number of observed variables for a given (hidden) state?

2. **Natural language processing** is one common application for HMMs: we have a single observation (a "word") that varies over time (a "sentence" or "document"), where each observation is associated with some property (like "part of speech").

   Consider the following HMM: $\prod[J, N, V]$ = [0.3, 0.4, 0.3]

A	J (adj)	N (noun)	V (verb)
J	0.4	0.5	0.1
N	0.1	0.4	0.5
V	0.4	0.5	0.1

B	*brown*	*leaves*	*turn*
J	0.8	0.1	0.1
N	0.3	0.4	0.3
V	0.1	0.3	0.6

   (a) How might we go about obtaining the values in the matrices Π, A, and B given above, in a **supervised** context?

   (b) Use the **forward** algorithm to find the probability of the "sentence" *brown leaves turn*.

   (c) Use the **Viterbi** algorithm to find the most likely state sequence for the sentence *brown leaves turn*.