





MITx 6.86x

Machine Learning with Python-From Linear Models to Deep Learning[Course](#)[Progress](#)[Dates](#)[Discussion](#)[Resources](#)[Course](#) / [Unit 3. Neural networks \(2.5 wee...](#) / [Lecture 11. Recurrent Neural Ne](#)[< Previous](#)

3. Markov Models to Feedforward Neural Nets

[Bookmark this page](#)

Exercises due Mar 29, 2023 08:59 -03 Past due

Feature Based Markov Models and Temporal/Sequence Problems**Video** [Download video file](#)**Transcripts** [Download SubRip \(.srt\) file](#) [Download Text \(.txt\) file](#)**Markov Transitions**

2 points possible (graded)

Suppose we represent a Markov model as a feedforward neural network, as described word, let the probability that word j occurs next be p_j . Which of the condition(s) below be the set of words. (Choose all that apply.)

☐ $\sum_{k \in K} p_k = 1$

☐ p_k is greater than or equal to zero for all $k \in K$ ☐ p_k is less than 0.5 for all $k \in K$

1 point possible (graded)

When representing a first-order Markov model as a feedforward network, what is the number of values in a single input vector?

☐ 0

☐ 1

☐ 2

☐ 3

Submit

You have used 0 of 2 attempts

Markov vs Feedforward

3 points possible (graded)

What are some advantages of the feedforward NN as described in the lecture versus Markov Models (select all that apply.)

☐ They contain a fewer number of parameters

☐ We can easily control the complexity of feedforward NN by introducing hidden layers

☐ They are able to encode more complex transition probabilities than Markov Models

Suppose you have a word vocabulary of size 10 (including <beg> and <end>), and you use a language model to predict the next word.

How many parameters would you need for a Markov Model?

☐ 1100

☐ 1001

☐ 1110

[Submit](#)

You have used 0 of 2 attempts

Discussion

Topic: Unit 3. Neural networks (2.5 weeks):Lecture 11. Recurrent Neural Networks 2 /
3. Markov Models to Feedforward Neural Nets

[Show all posts](#) ▼

? What are parameters in the Markov Model

In Markov Model, instead of using a neural network approach we're using a probabilistic approach , right? So

? PKs

The Professor said that the transformation we need is that "PKs are non-zero". But what he wrote is $PK \geq 0$

? How are the first and last words handled in a trigram model when using a Markov process?

I'm starting with a 4 word trigram model to get the formula for the number of parameters needed by the diffe

💬 Can all information, which is possible to encode in a Markov model, be encoded in a neural ne

Any information encoded in a neural network could also be encoded in a very large transition probability mat

💬 1st and 2nd order Markov models

The transformation from a 1st order Markov model to a FFNN is deemed a perfectly valid transformation but

edX

[About](#)[Affiliates](#)[edX for Business](#)[Open edX](#)[Careers](#)[News](#)

Legal

[Terms of Service & Honor Code](#)[Privacy Policy](#)



© 2023 edX LLC. All rights reserved.

深圳市恒宇博科技有限公司 [粤ICP备17044299号-2](#)