Feed-forward neural network design

For each application, describe a neural network structurally suited to solving the problem.

Your descriptions should include at least the following elements:

- Encoding for the inputs
  How are the input predictors mapped to a fixed-size sequence of real numbers?
- Sequence of layers
  You may define each layer as a weight matrix (how big?), optional bias vector (do you want it?), and activation function (which one?).
- Loss function

Linear regression

Predict $\hat{y}$ from $\bar{x}$, modeling the phenomenon as $y = \bar{w}^T \bar{x}$. y is a scalar and $\bar{x}$ has length d.

Classification

Given the height and weight of an NFL player, predict what position they play: wide receiver, running back, or quarterback (ignore other positions).

Bigram language modeling

Given a token, predict the [distribution of the] next token. The vocabulary size is V.

Trigram language modeling

Given two (ordered) tokens, predict the [distribution of the] next token. The vocabulary size is V.

Trigram language modeling, but better

Imagine that your model from the previous part does not produce sufficiently accurate results. You suspect underfitting. Design a more expressive model for the same application, and explain why you designed it this way.