# Neural Language Models

**Phạm Quang Nhật Minh**
AImesoft JSC
minhpham0902@gmail.com

February 11, 2023

# Table of Contents

# Neural Language Models (LMs)



- **Language Modeling**: Calculating the probability of the next word in a sequence given some history.
  - We've seen N-gram based LMs
  - But neural network LMs far outperform n-gram language models
- State-of-the-art neural LMs are based on more powerful neural network technology like Transformers
- But **simple feedforward LMs** can do almost as well!

# Simple feedforward Neural Language Models



- **Task**: predict next word $w_t$

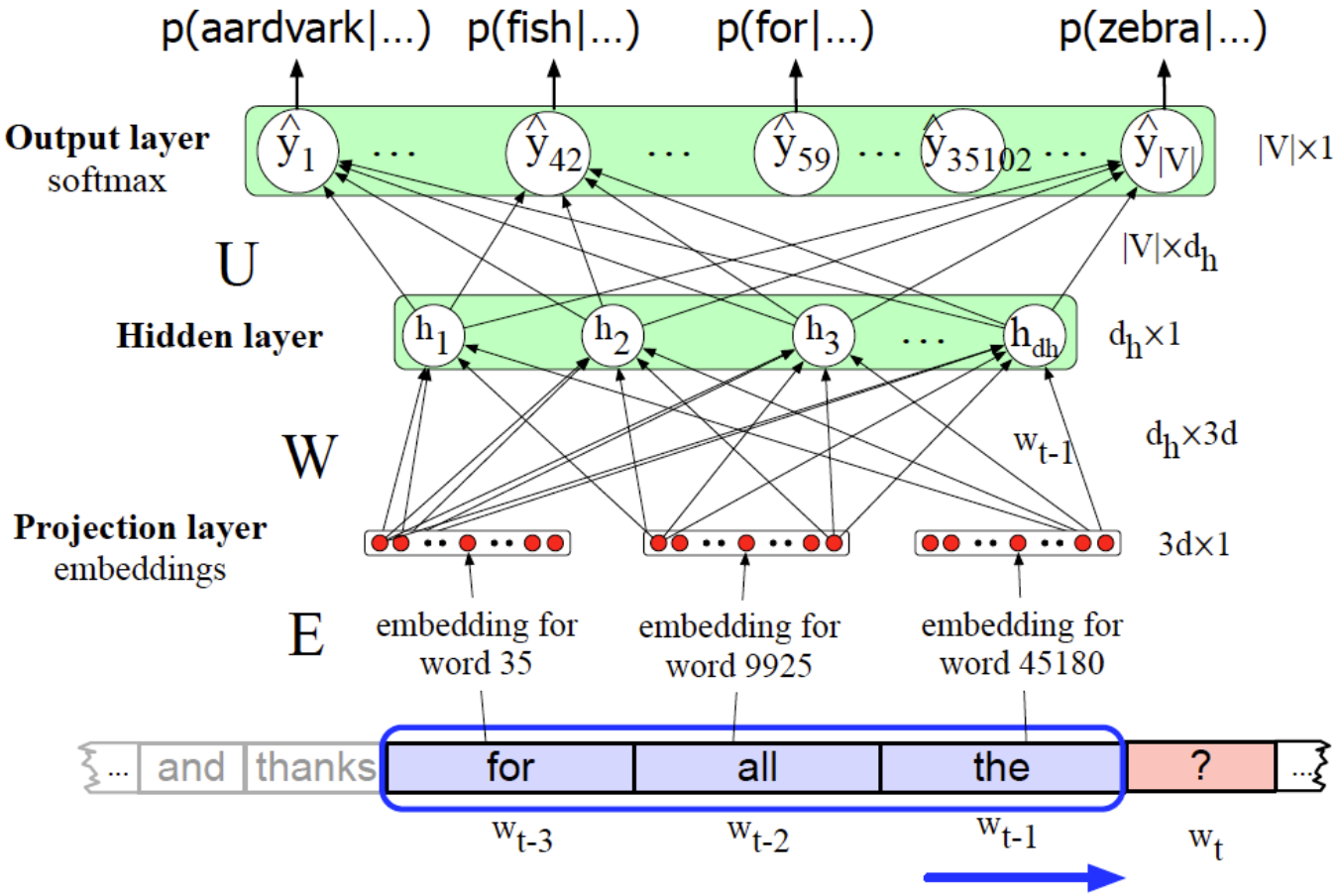
  given prior words $w_{t-1}$, $w_{t-2}$, $w_{t-3}$, ...

- **Problem**: Now we're dealing with sequences of arbitrary length.
- **Solution**: Sliding windows (of fixed length)

$$P(w_t|w_1^{t-1}) \approx P(w_t|w_{t-N+1}^{t-1})$$

# Neural Language Model





# Why Neural LMs work better than N-gram LMs



- **Training data:**
- We've seen: I have to make sure that the cat gets fed.
- Never seen: dog gets fed
- **Test data:**
- I forgot to make sure that the dog gets ___
- N-gram LM can't predict "fed"!
- Neural LM can use similarity of "cat" and "dog" embeddings to generalize and predict “fed” after dog

# Training neural language models