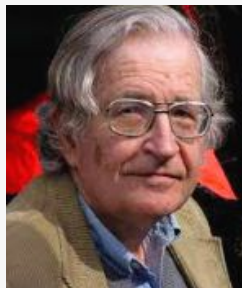# Module 5

NLP History

Noam Chomsky (1950s-1960s):
- Chomsky's work on transformational-generative grammar laid the foundation for modern linguistics and NLP.
- He proposed the idea of a "universal grammar" that underlies all human languages.

Yoshua Bengio (2000s):
- Bengio's research focused on neural networks and deep learning, which became essential in the development of modern NLP models.
- He contributed to the development of word embeddings, representing words as dense vectors in continuous vector spaces.

Tomas Mikolov (2010s):
- Mikolov's research at Google led to the development of Word2Vec, a popular word embedding technique.
- Word2Vec efficiently represented words as numerical vectors, capturing semantic relationships between them.

NLP History

Transformer Model (2017):

- The Transformer model was introduced in the paper "Attention Is All You Need" by Vaswani et al.
- It utilized the self-attention mechanism, allowing the model to weigh the importance of different words in a sentence when processing each word.
- Transformers proved to be highly effective for various NLP tasks and outperformed traditional sequence-to-sequence models in machine translation.

**GPT (Generative Pre-trained Transformer) (2018):**

- GPT, developed by OpenAI, was one of the first large-scale language models based on the Transformer architecture.
- It was pre-trained on a massive corpus of text data and fine-tuned for specific NLP tasks.
- GPT demonstrated remarkable language generation capabilities and achieved state-of-the-art performance on various language tasks.

Word embeddings

Performing math operations on words: https://turbomaze.github.io/word2vecjson/

Vectors Continued...
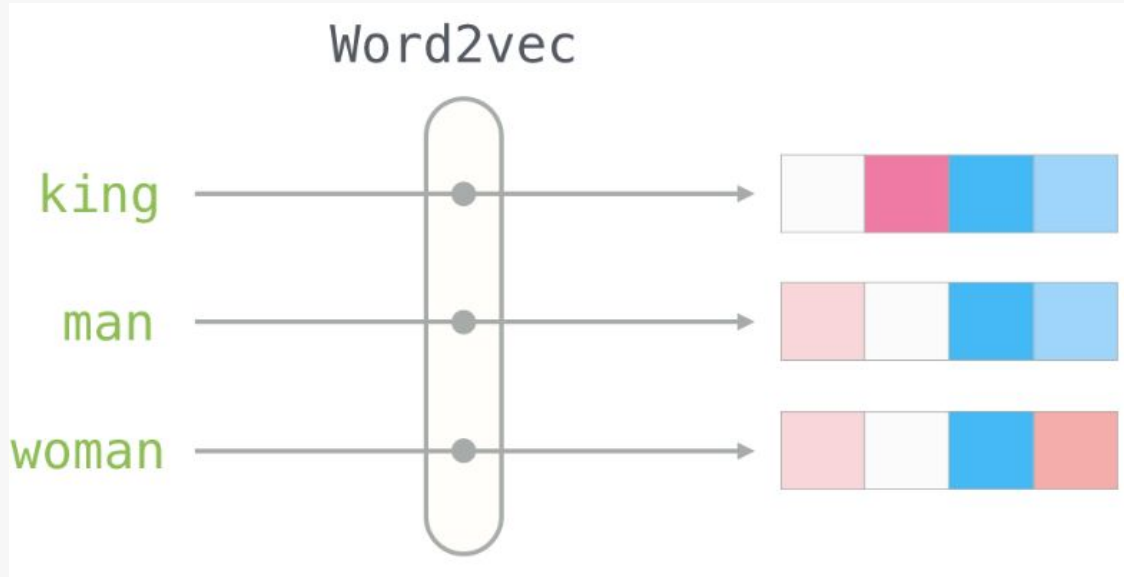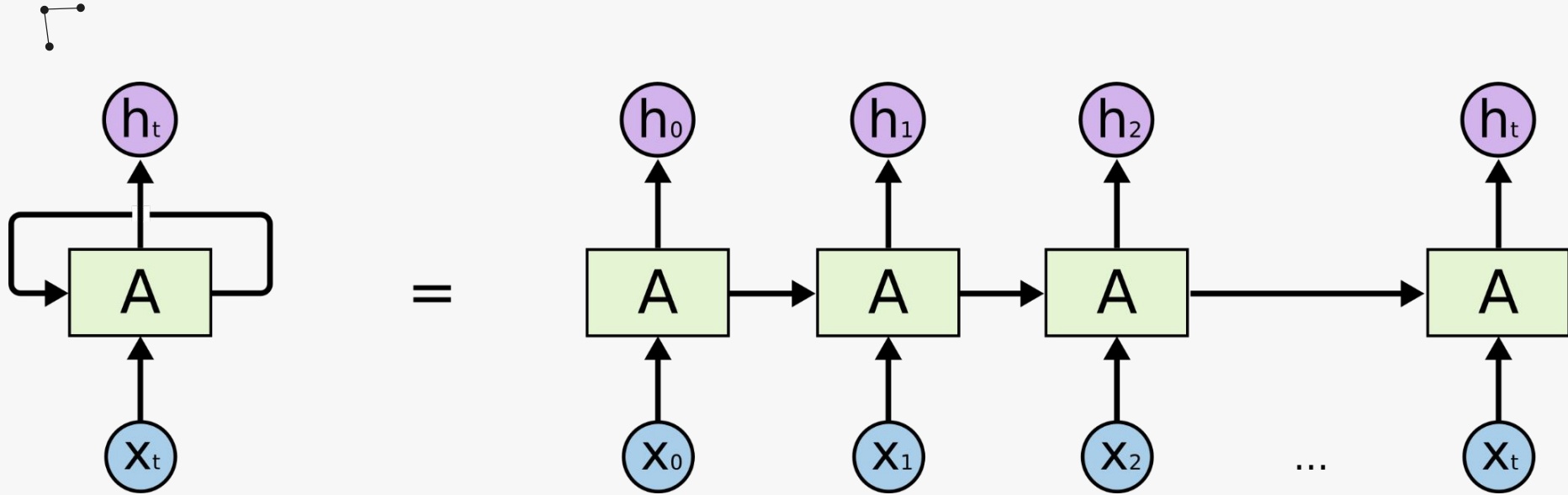
Source: Jay Alammar's <u>blog post</u>

Natural Language Processing

- Basic models - TFIDF
- NLTK processing - Lemmatization, Stemming, rule based processing (hindi/kannada/english rule)
- Tokenization
- Word vectors - embeddings
- RNNS
- LSTM and GRU

Recurrent Neural Networks

Loops in Neural Networks & Autoregressive Models