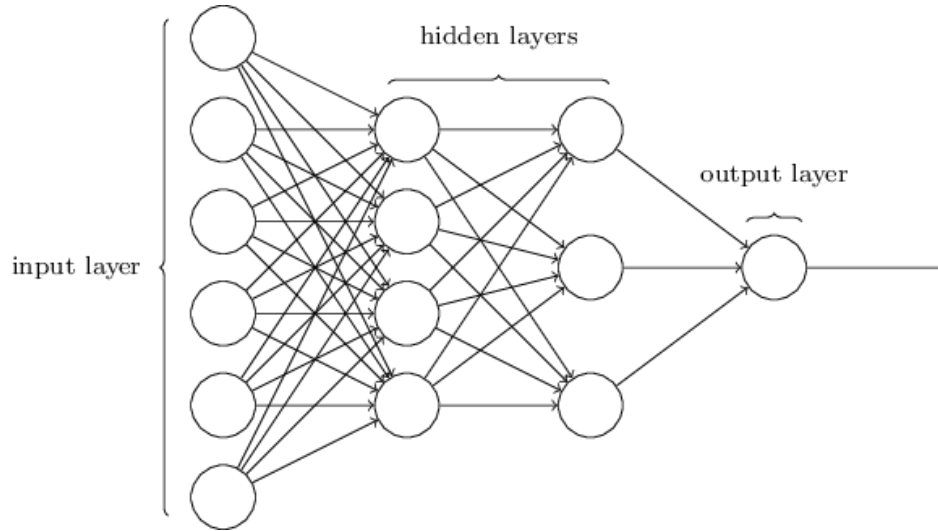# Dynamic Convolutional Network for Sentence Modeling

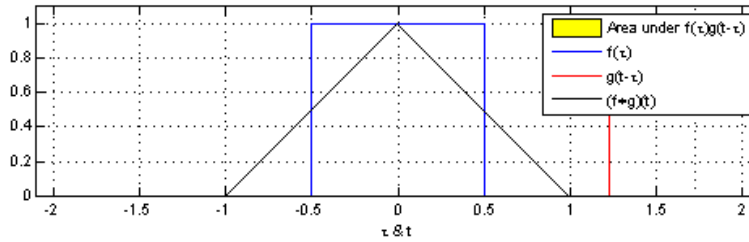## CS365 - Midterm Presentation

Group 13

Jayesh Kumar Gupta (11337) & Arpit Shrivastava (12161)
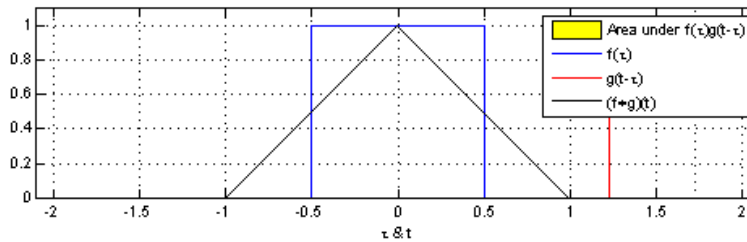
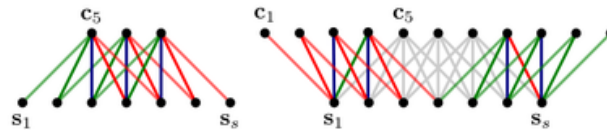# What is a Neural Network?

What is Convolution?

$$(f * g)(t) = \sum_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$

What is Convolution?

$$(f * g)(t) = \sum_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$

Types

Narrow and Wide Types of Convolution

What is Convolutional Neural Network (CNN)?

What is Convolutional Neural Network (CNN)?

- Can be seen as a kind of neural network that uses many identical copies of the same neuron.
- Can express computationally large models with lesser number of parameters.
- Each neuron takes inputs from a rectangular section of the previous layer.
- The weights for this rectangular section are the same for each neuron in the convolutional layer.
- These weights act as the convolution filter.

What is Convolutional Neural Network (CNN)?

Structure
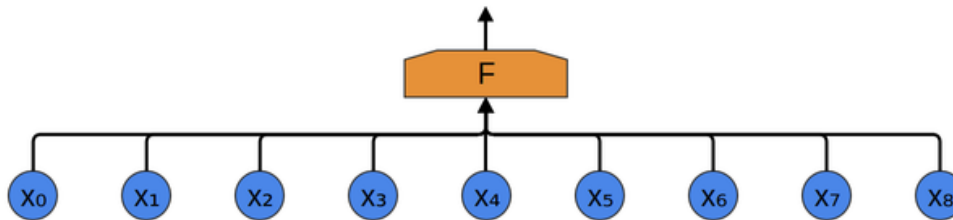
- Say you have data samples as:

$$X_0 \quad X_1 \quad X_2 \quad X_3 \quad X_4 \quad X_5 \quad X_6 \quad X_7 \quad X_8$$

What is Convolutional Neural Network (CNN)?

Structure

- Say you have data samples as:

X_0 X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8

- The simplest way to classify them would be to just connect them all to a fully-connected layer where every input connects to every neuron.

F

X_0 X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8
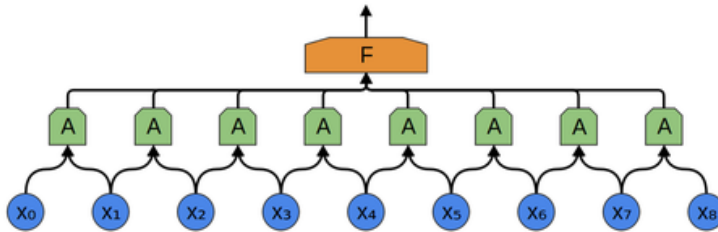
# What is Convolutional Neural Network (CNN)?

Structure

- But there are a lot of local properties and symmetries that can be useful. Hence we create a group of neurons to look at only a segment of data and compute certain *features*. This is the *convolutional layer*.

- These layers can be composed together to learn higher level abstract features from the data.
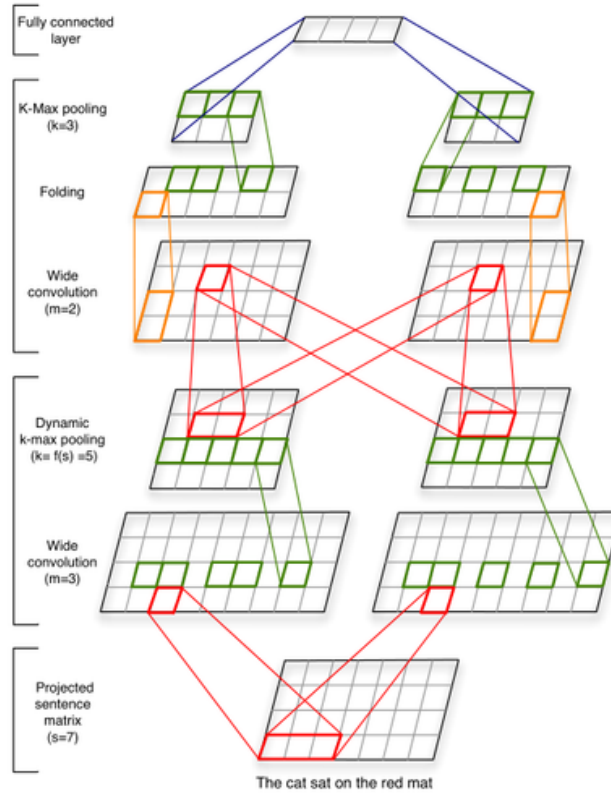
What is Convolutional Neural Network (CNN)?

Pooling
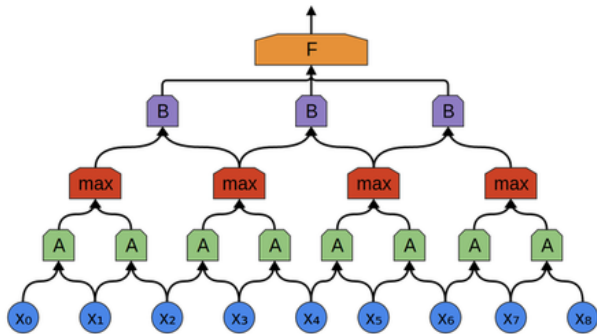
- Convolutional layers are often interweaved with pooling layers.
- A popular approach is **max-pooling** where we take the maximum of features over small blocks of a previous layer, making the network invariant of small transformations in the data.

What is Dynamic CNN?

What is Dynamic CNN?

k-Max Pooling

- Generalization of max-pooling operation.
- Instead of selecting a single max value, we select a subsequence of length k of max values.

- With l as current layer, L as total number of layers and k_{top} as pooling parameter for topmost layer, pooling parameter, k is modeled as:

$$k(l) = \max(k_{top}, \left\lceil \frac{L-l}{L} s \right\rceil)$$

How is the sentence modeled?

How is the sentence modeled?

Word and n-Gram Order

- Filters m of wide convolution in the first layer learn to recognize specific n-grams that have size less than or equal to the filter width m.
- Pooling induces invariance to absolute positions of these subsequence of n-grams while maintaining their order and relative positions.

How is the sentence modeled?

Word and n-Gram Order

- Filters m of wide convolution in the first layer learn to recognize specific n-grams that have size less than or equal to the filter width m.
- Pooling induces invariance to absolute positions of these subsequence of n-grams while maintaining their order and relative positions.

Induced Feature Graph

- Layers of convolution and pooling induce a directed acyclic graph with weighted edges over the input called the induced feature graph.
- The dynamic k-max pooling operation allows the network to draw together features that correspond to words far apart in a sentence.

Questions?

Thank You!

Slides were made with remark.js

All images stolen blatantly from the internet.