Deep Knowledge-Aware Network (DKN) for News Recommendation

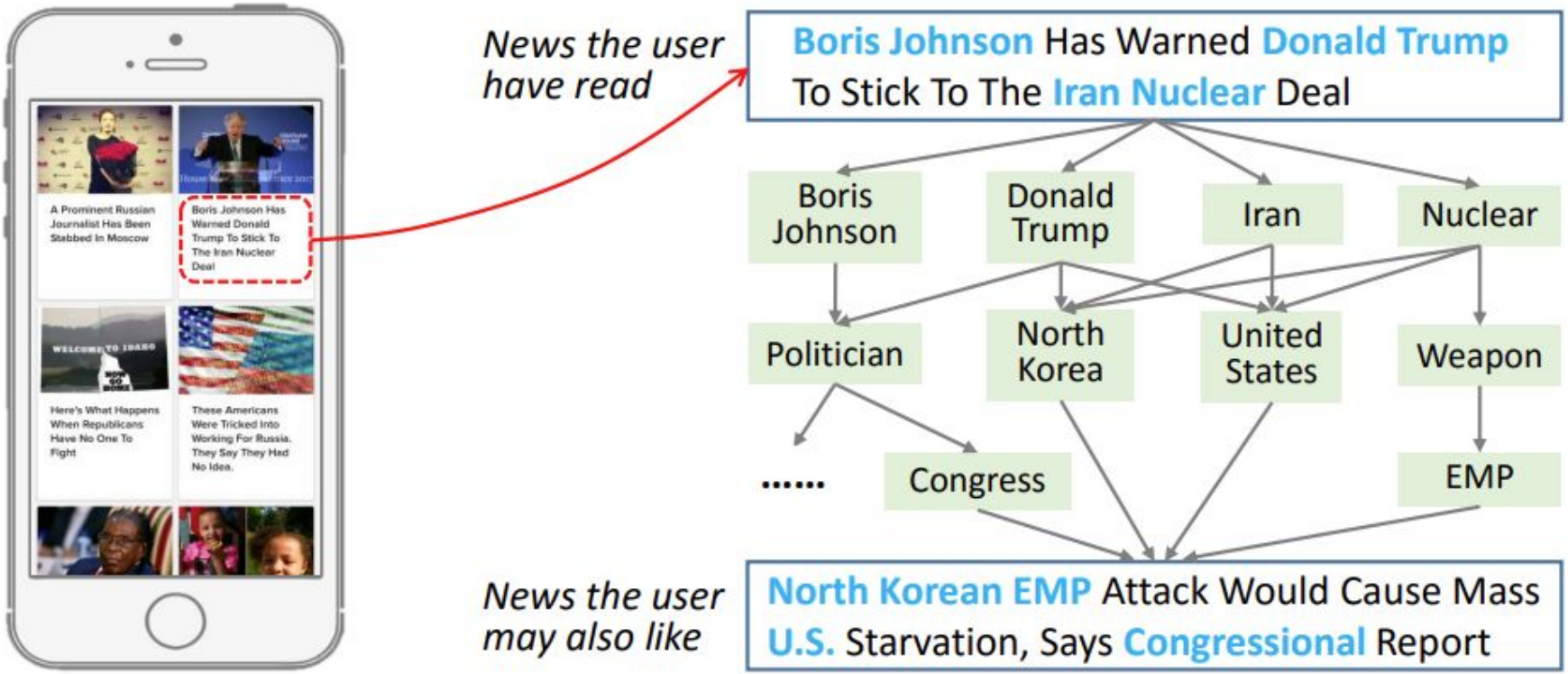
Introduction

- Change in News consumption habits
- A challenge of online news platforms is the volume of articles
- Help users by making personalized recommendations

Challenges in News Recommendation

- News articles are highly time-sensitive - This makes methods such as collaborative filtering less effective
- People are topic-sensitive
- News language is highly condensed and comprised of knowledge entities and common sense

News Connected through Knowledge Entities



Objective of DKN

It takes candidate news and one user’s click history as input and outputs the probability of the user clicking the news.

Properties of DKN

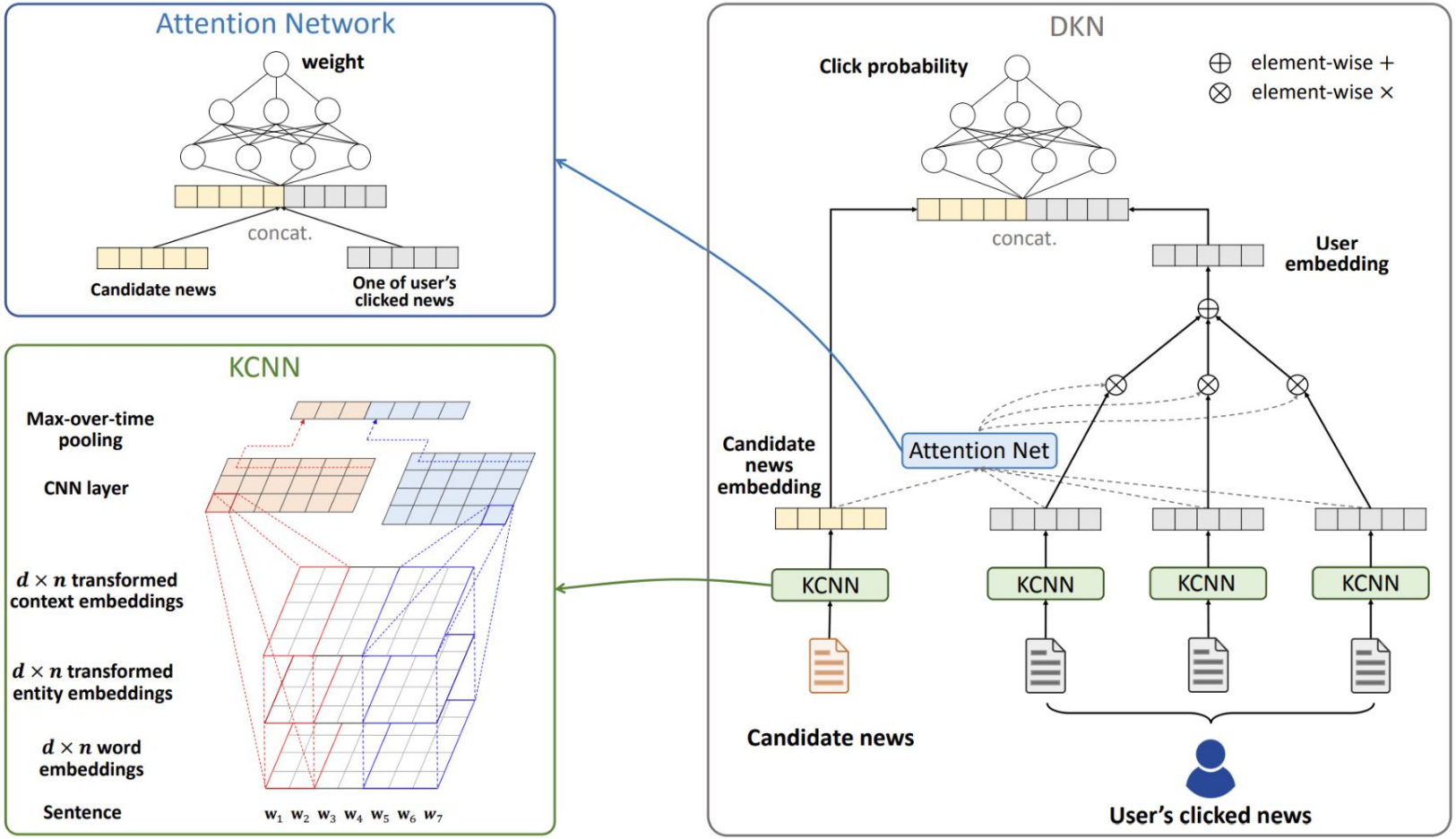
Content based Deep Learning model for CTR prediction

It makes use of Knowledge entities and news content for recommendations

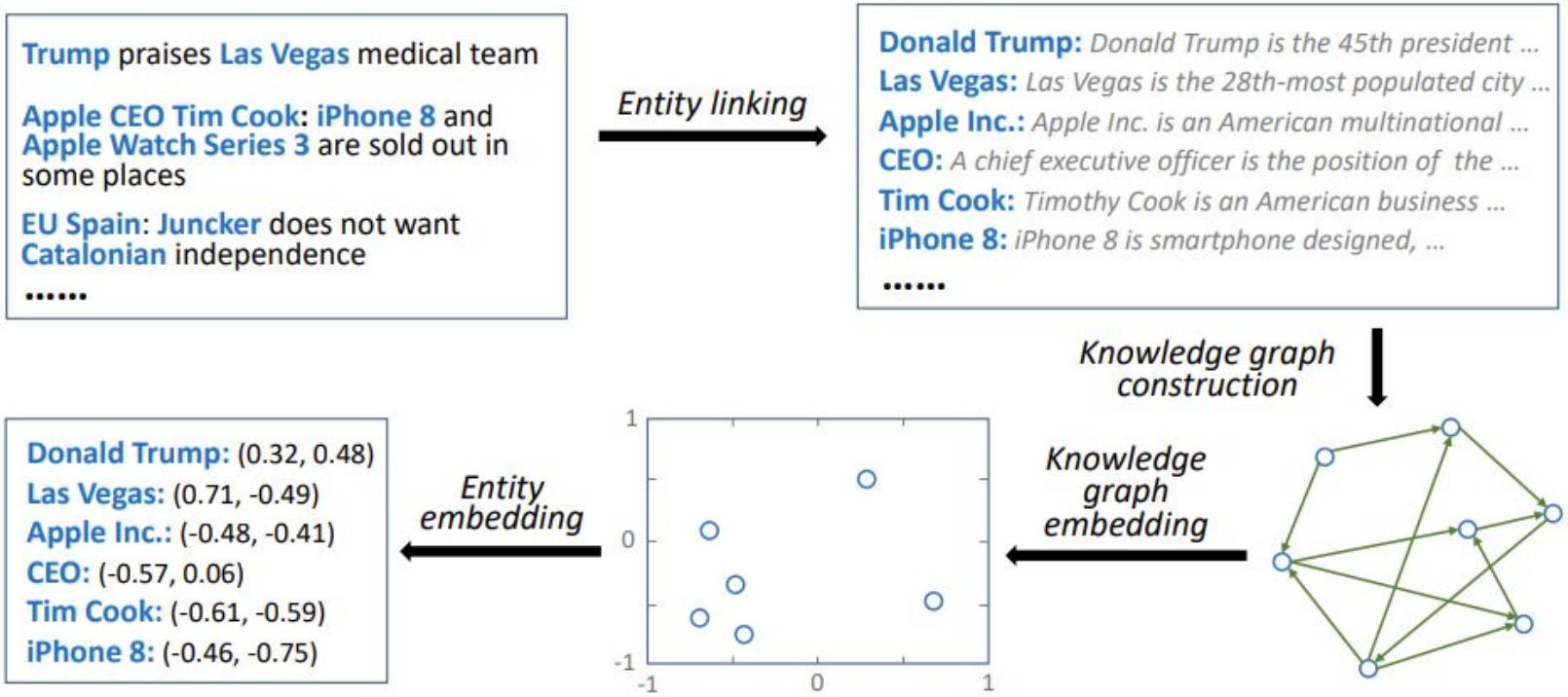
Knowledge-aware convolutional neural networks (KCNN) are used to fuse word-level and knowledge-level representations of news and generate a knowledge-aware embedding layer

It uses attention module to calculate a user's aggregated historical representation

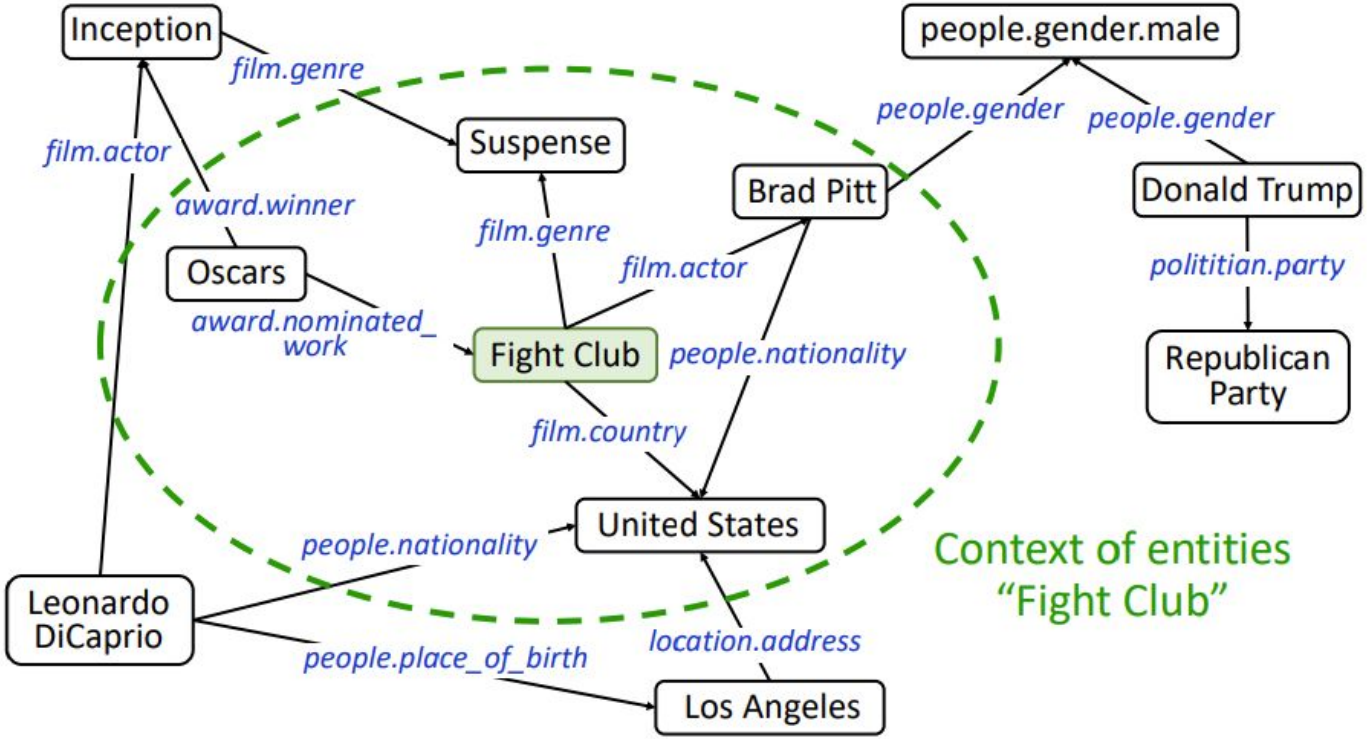
DKN Architecture



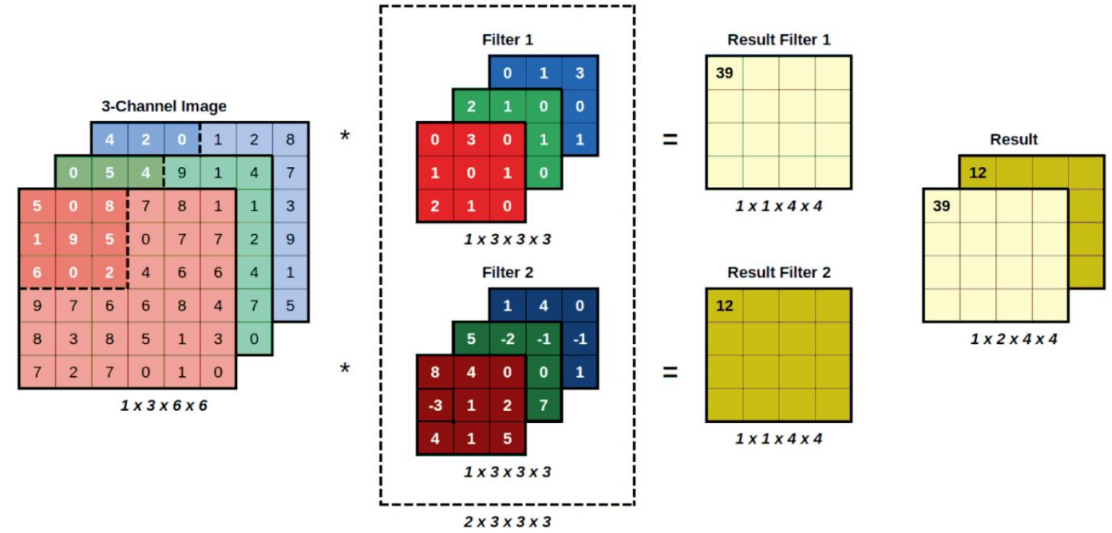
Knowledge Distillation

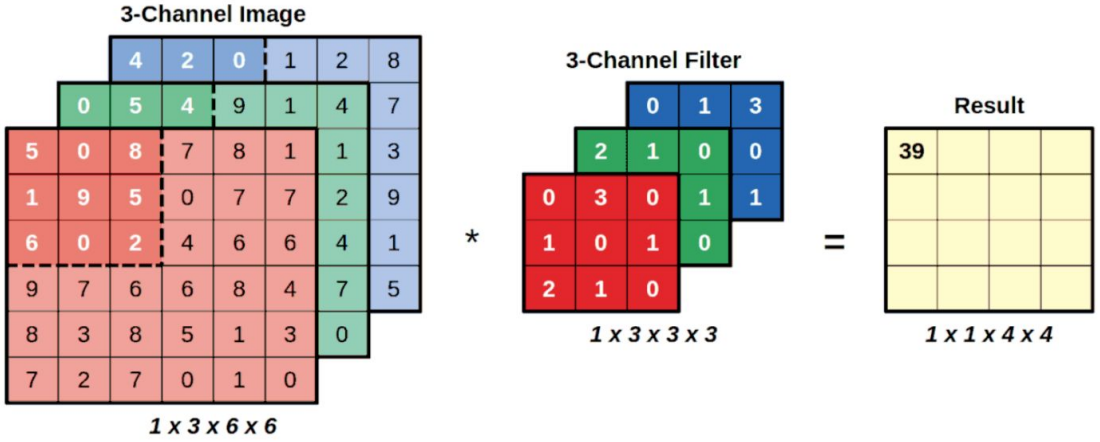


Context of an Entity in a Knowledge Graph

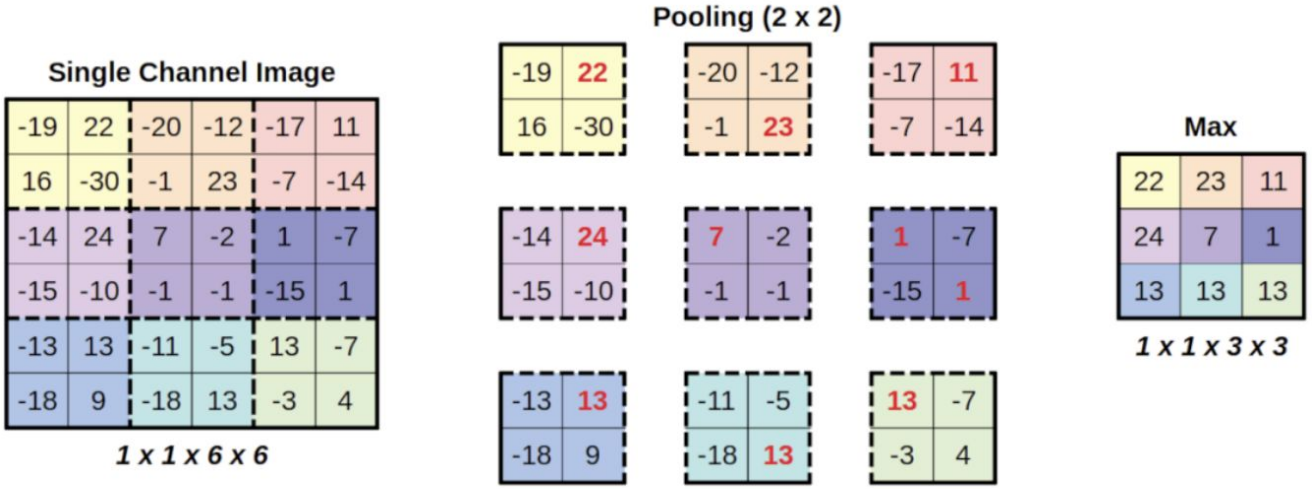


How Convolution works

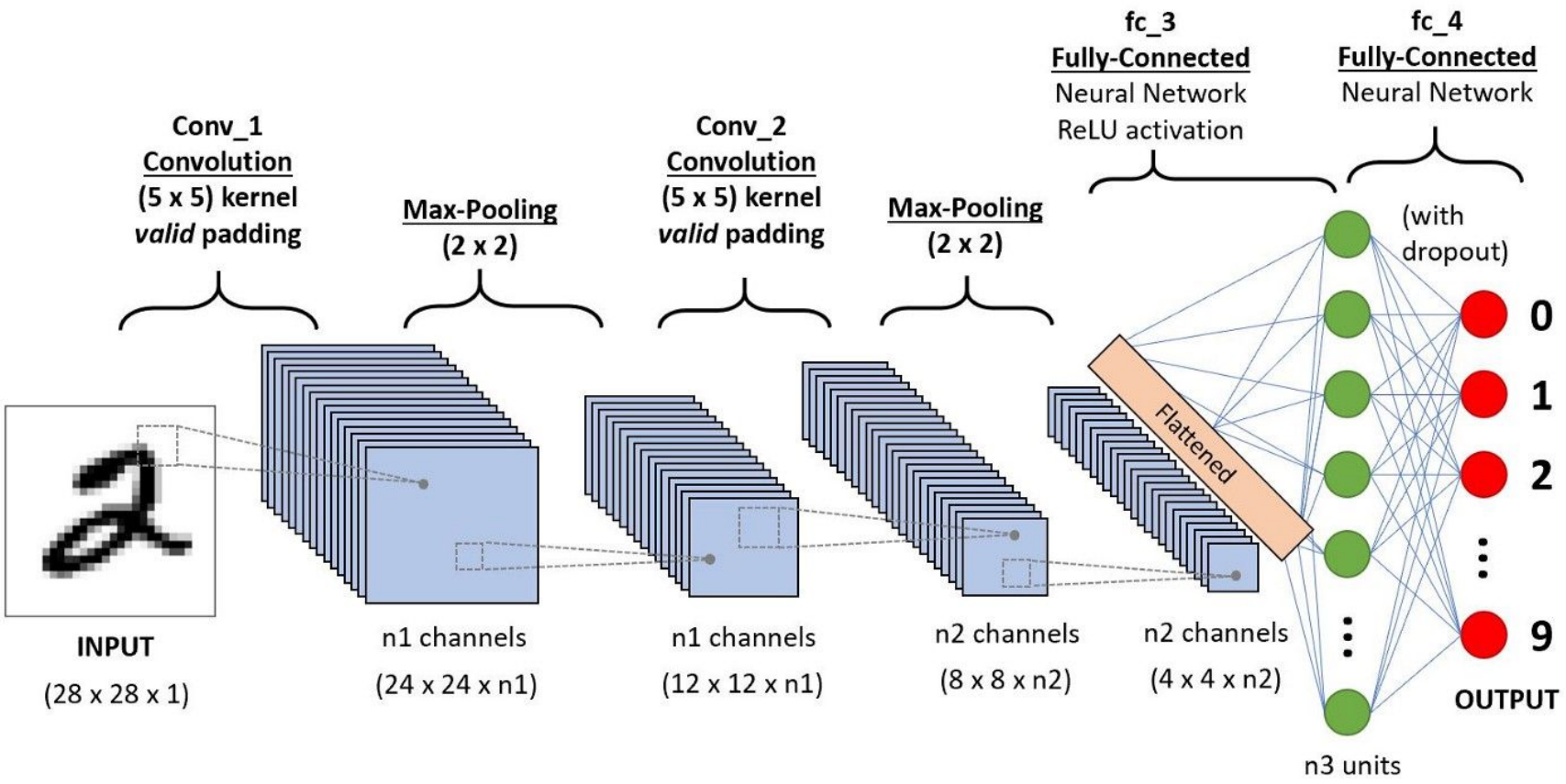




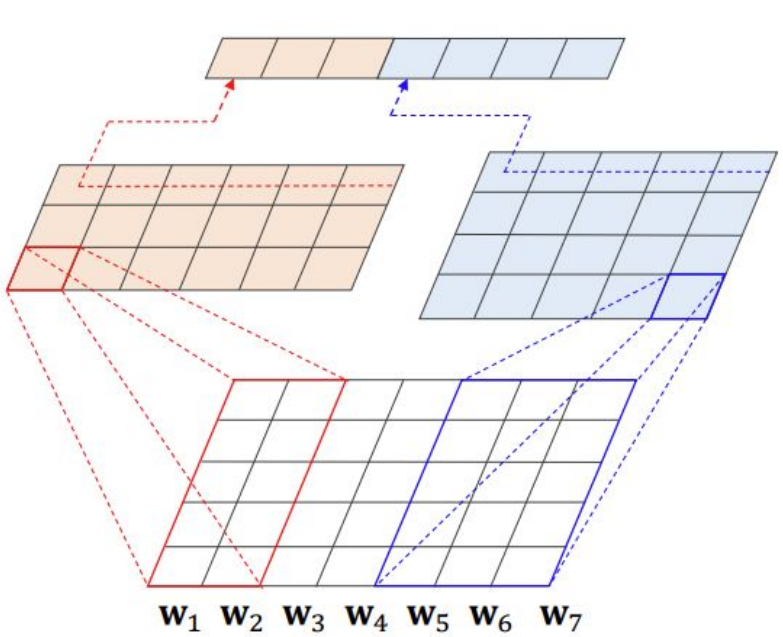
Max Pooling


Single Channel Image
-19 22 -20 -12 -17 11
16 -30 -1 23 -7 -14
-14 24 7 -2 1 -7
-15 -10 -1 -1 -15 1
-13 13 -11 -5 13 -7
-18 9 -18 13 -3 4
1 x 1 x 6 x 6
Pooling (2 x 2)
-19 22
16 -30
-20 -12
-1 23
-17 11
-7 -14
-14 24
-15 -10
7 -2
-1 -1
1 -7
-15 1
-13 13
-18 9
-11 -5
-18 13
13 -7
-3 4
Max
22 23 11
24 7 1
13 13 13
1 x 1 x 3 x 3

CNN


Conv_1
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
Conv_2
Convolution
(5 x 5) kernel
valid padding
Max-Pooling
(2 x 2)
fc_3
Fully-Connected
Neural Network
ReLU activation
fc_4
Fully-Connected
Neural Network
(with dropout)
Flattened
INPUT
(28 x 28 x 1)
n1 channels
(24 x 24 x n1)
n1 channels
(12 x 12 x n1)
n2 channels
(8 x 8 x n2)
n2 channels
(4 x 4 x n2)
n3 units
0
1
2
9
OUTPUT

CNN for Sentence Representation learning



$w_{1:n} = [Donald\ Trump\ praises\ Las\ Vegas\ medical\ team]$

Sentence representation

Max-over-time pooling

Feature maps

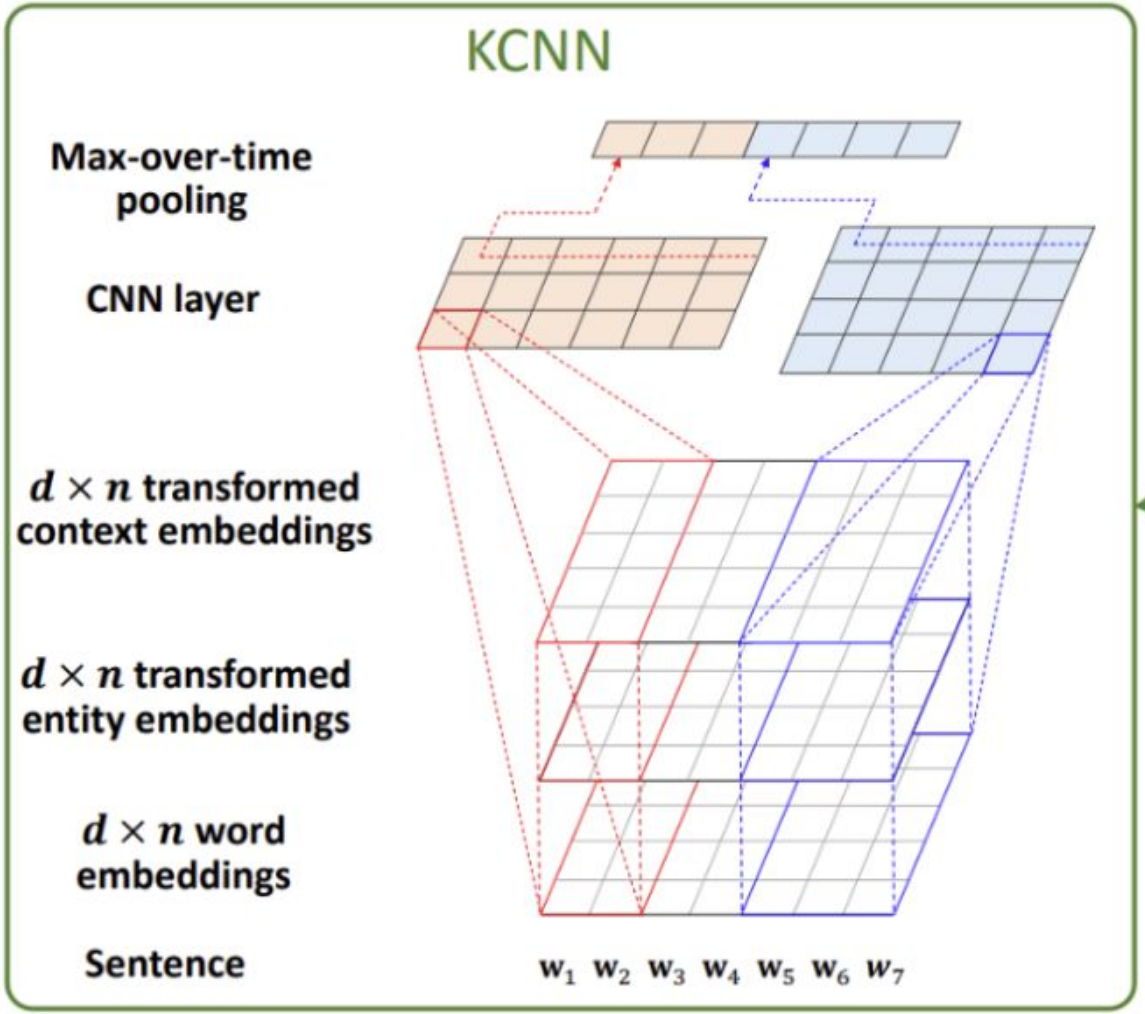
Convolution

$d \times n$ word embedding matrix

Sentence

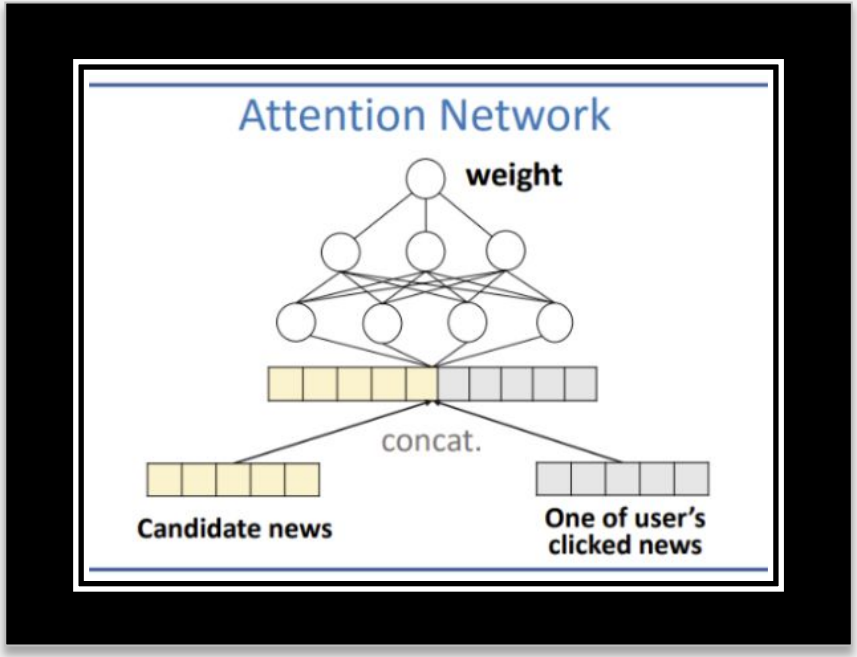
KCNN

- It is multi-channel
  - Word embedding
  - Entity embedding
  - Contextual entity embedding
- The above embeddings are stacked just like colour images
- KCNN provides knowledge-aware representation vector for each piece of news

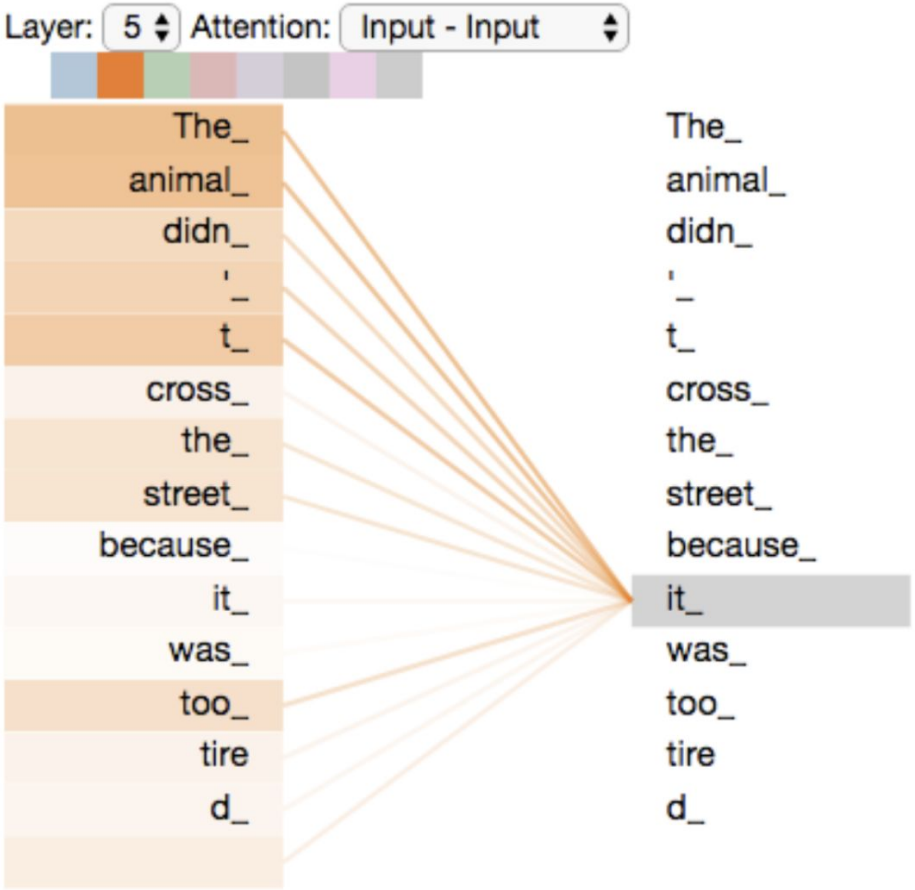


Attention Module

- To get dynamic representation of the user
- It matches candidate news to each piece of clicked news
- User’s history is aggregated with different weights

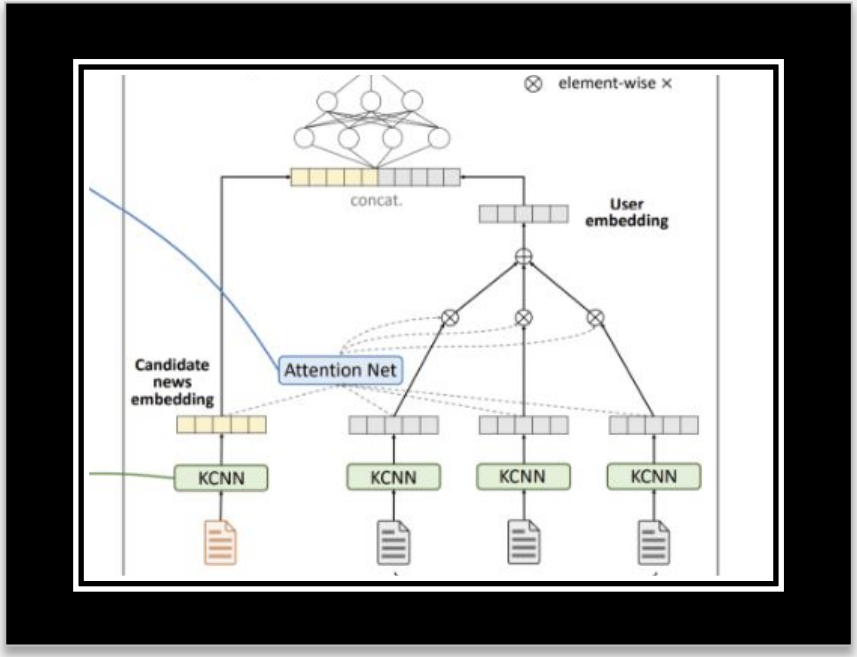


Attention


Layer: 5
Attention: Input - Input
The_
animal_
didn_
'_
t_
cross_
the_
street_
because_
it_
was_
too_
tire
d_
The_
animal_
didn_
'_
t_
cross_
the_
street_
because_
it_
was_
too_
tire
d_

Deep Neural Network (DNN)

- Candidate news embedding and user's embedding are finally processed by DNN for CTR prediction



Summary

- DKN takes advantage of Knowledge graph representation in news recommendation
- Content Based Deep learning model
- KCNN is used to jointly learn from semantic and knowledge level representations of news
- Attention module is used to calculate a user’s aggregated historical representation