Balancing between Holistic and Cumulative Sentiment Classification

Pantelis Agathangelou, Ioannis Katakis

University of Nicosia

Published in **Online Social Networks and Media** - Journal

Extends **:** *A Hybrid Deep Learning Network for Modelling Opinionated Content*

Summary

CUMULATIVE VS HOLISTIC SENTIMENT ASSIGNMENT

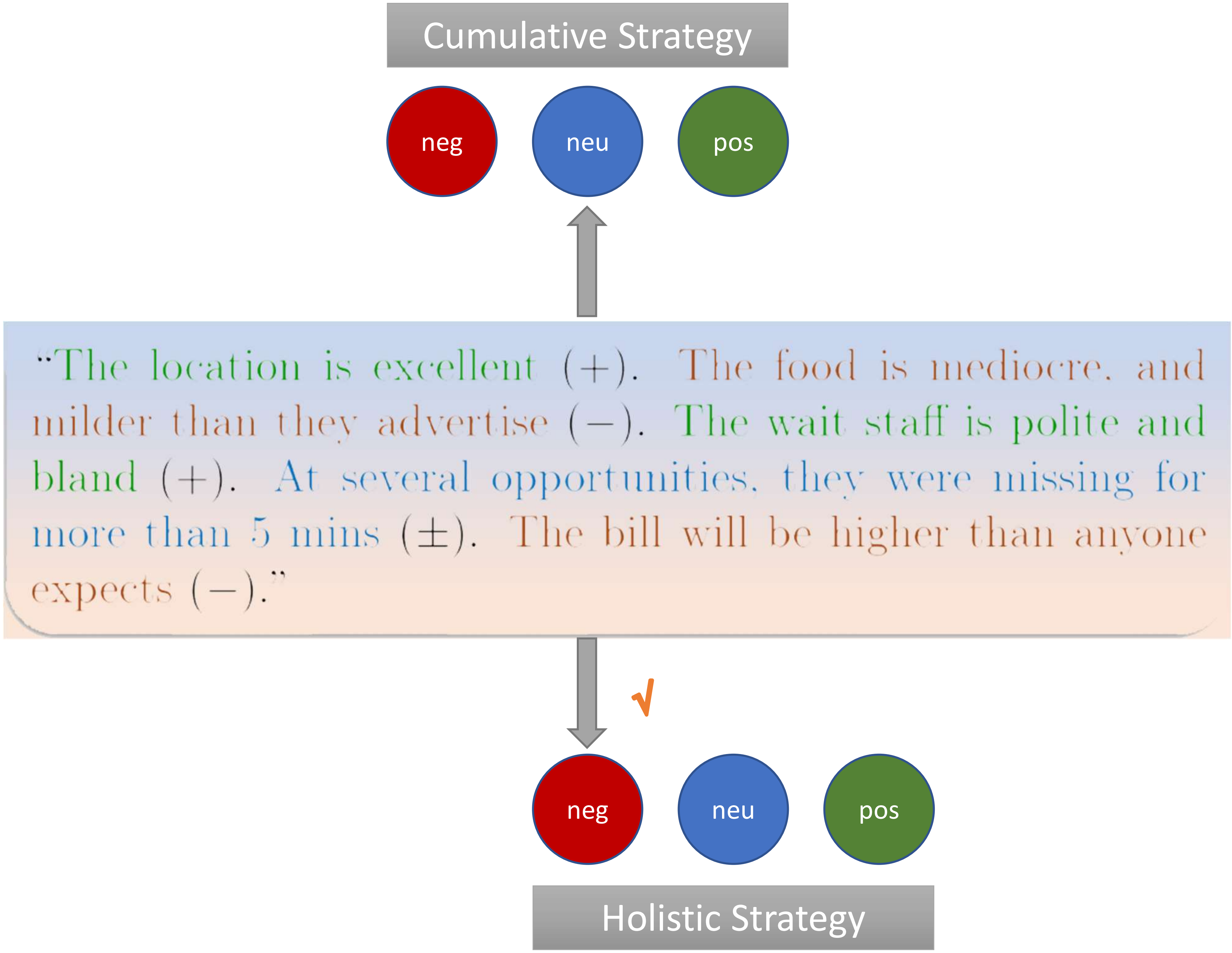
MODEL ARCHITECTURE

COMPONENT'S ANALYSIS

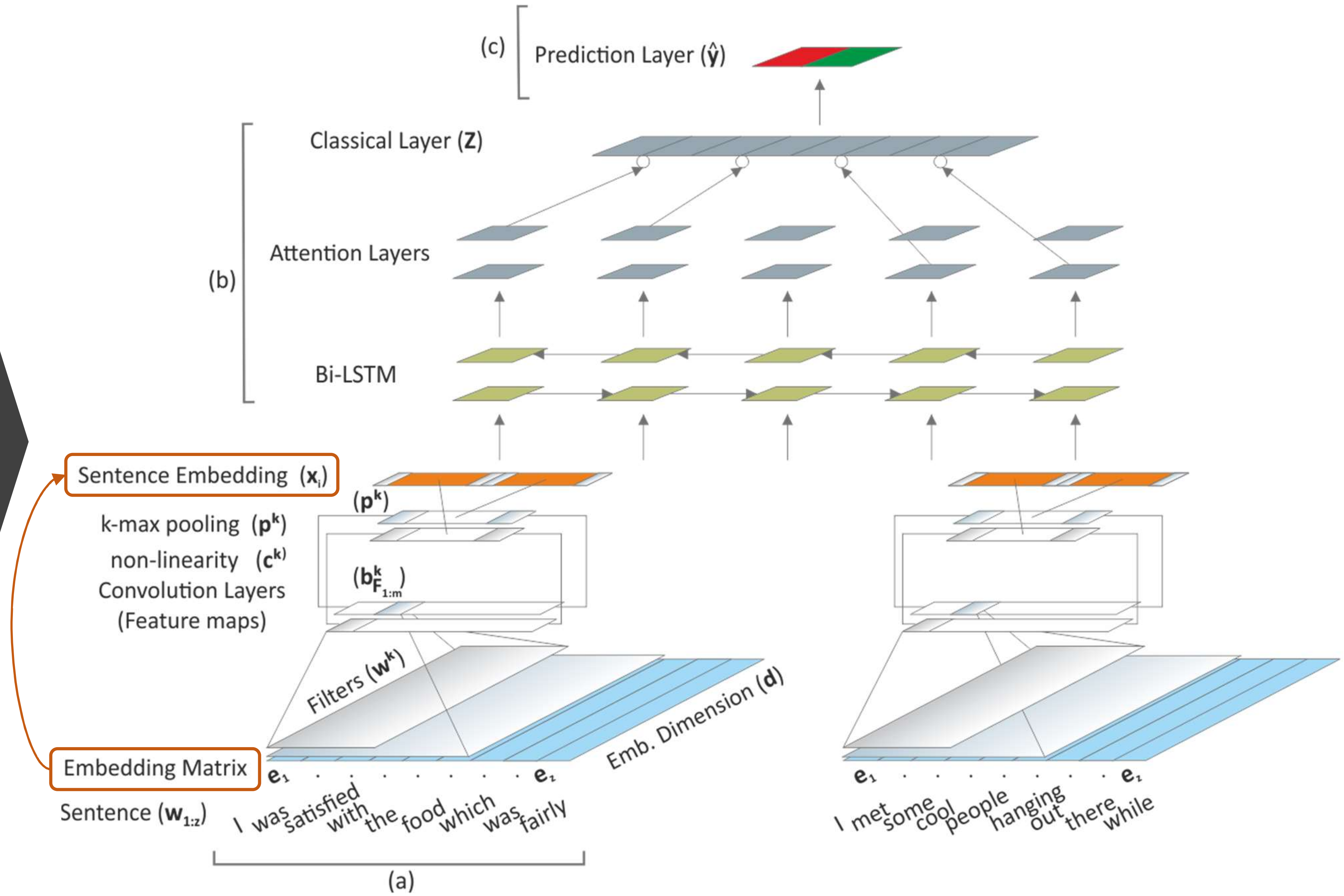
EXPERIMENTAL EVALUATION

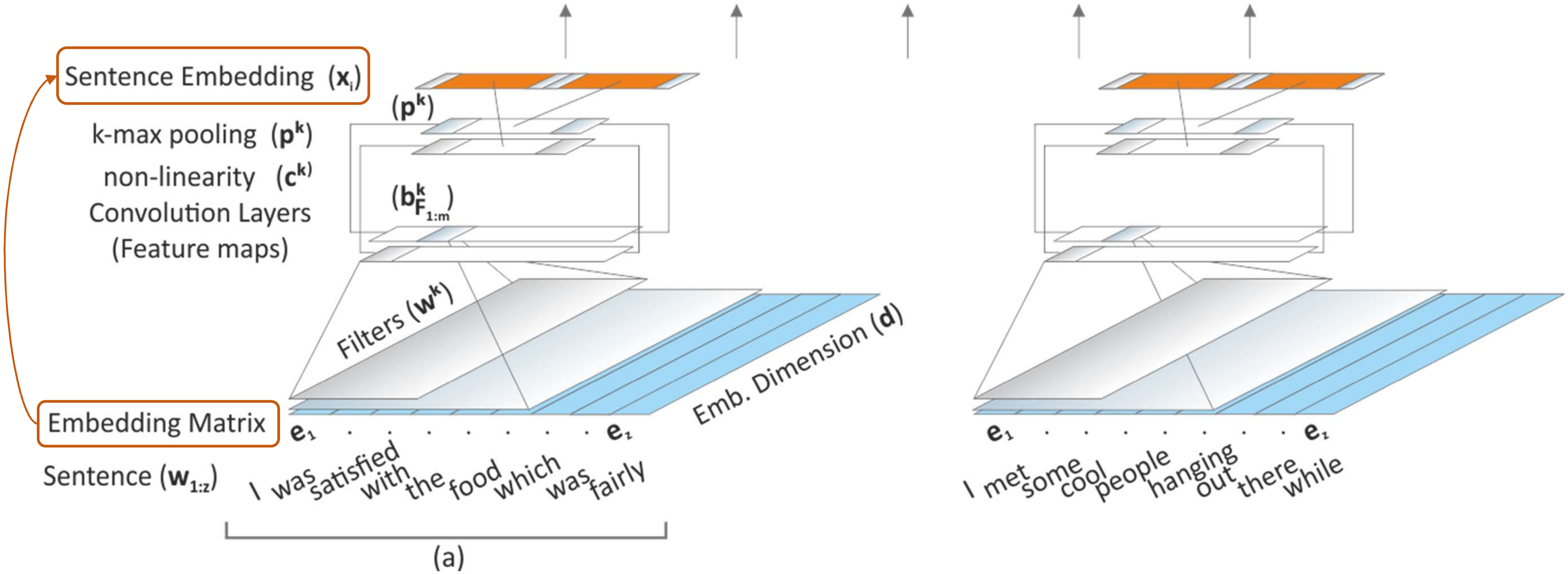
CONCLUSION

Cumulative vs Holistic Sentiment Assignment



Model Architecture



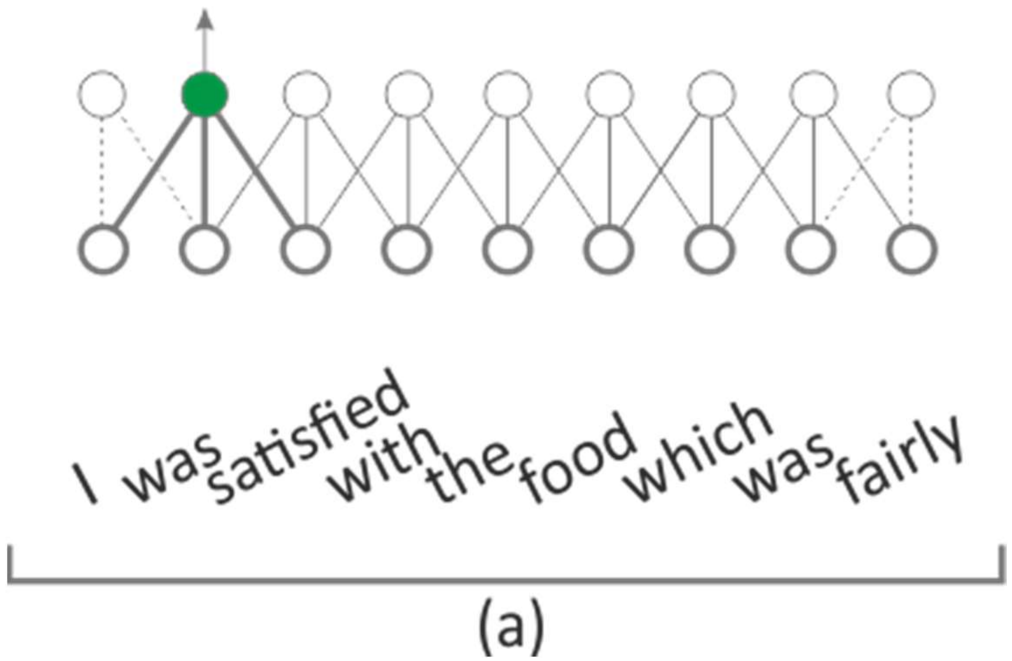


Exploring Sentiment Patterns

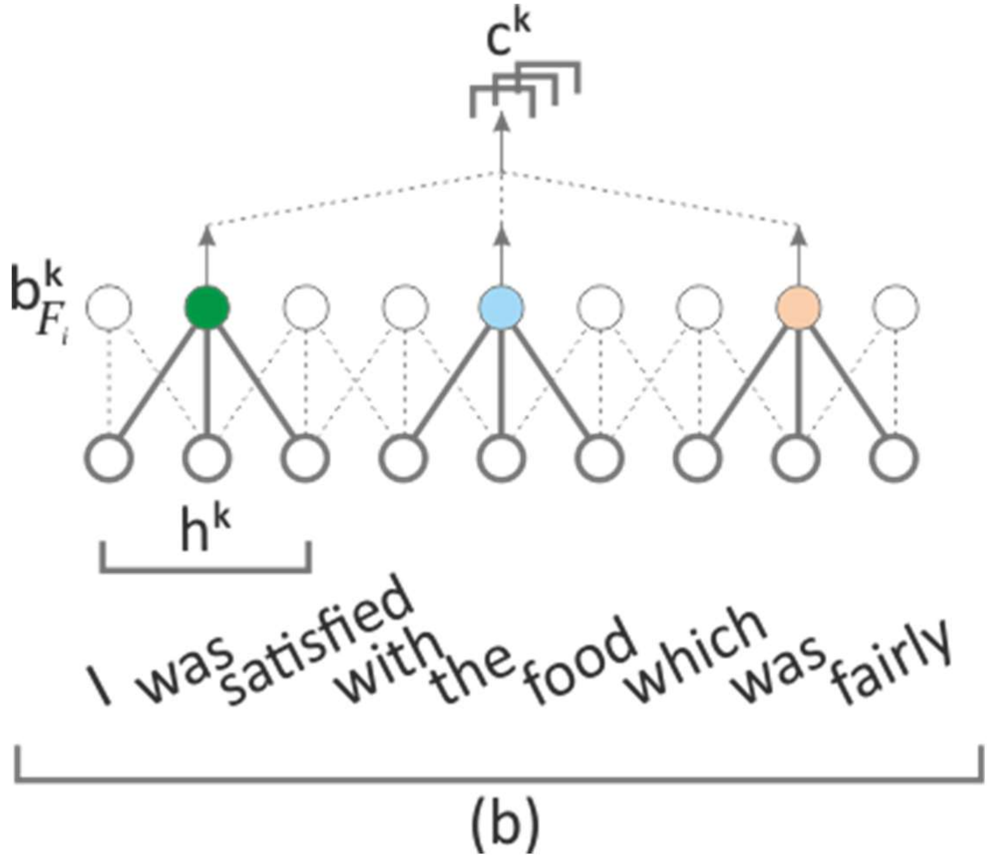
Vector Element	Feature Map Vector (length Z)
$b_i^k = w^k \cdot e_{i:i+h-1}$ (1)	$b_{F_i}^k = [b_1^k; \cdots; b_i^k; \cdots; b_z^k]$ (2)
Array of Feature Map Vectors	Non-Linearity
$b_{F_{1:m}}^k = [b_{F_1}^k; \cdots; b_{F_i}^k \cdots; b_{F_m}^k]$ (3)	$c^k = f^{nl}\left(b_{F_{1:m}}^k, \beta^k\right)$ (4)
k-max pooling	Sentence Embedding
$p^k = k\text{-}max[c^k]$ (5)	$x_i = [p^1; \cdots; p^k]$ (6)

single-max pooling VS k-max pooling in a Convolution Layer

Single max-pooling



K-max-pooling

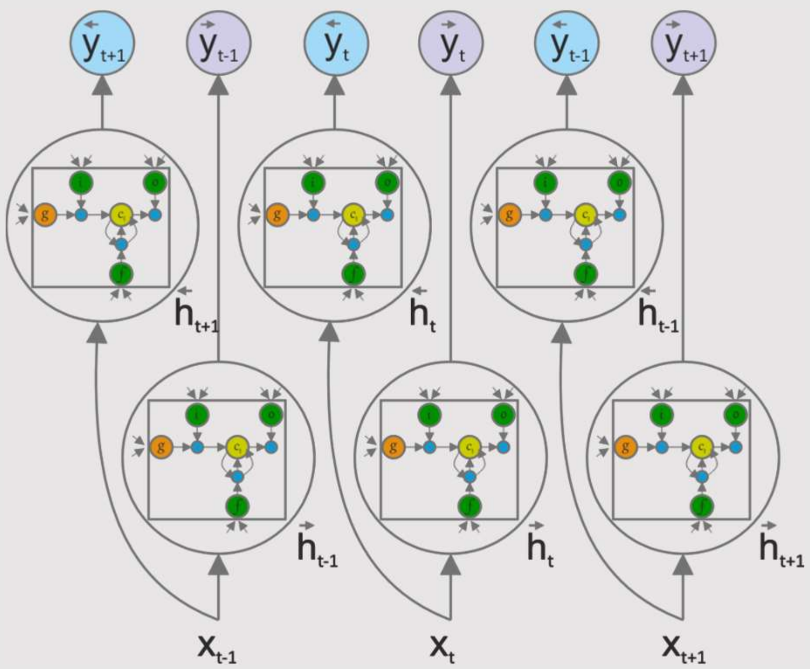


Pantelis Agathangelou, Balancing Between Holistic and Cumulative Sentiment Classification

Encoding Opinion – Augmenting Features – Grasping Sentiment Assignment Strategy

Recurrent Layer

(b)



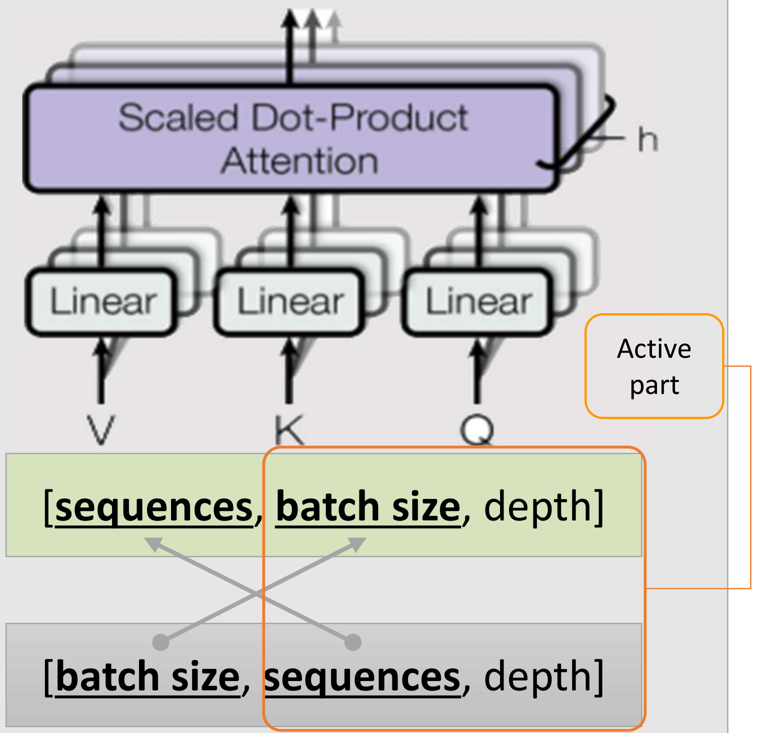
Forward Output Predictions

$$\overrightarrow{y_t} = W_{\overrightarrow{h}y} \cdot \overrightarrow{h}_t + \overrightarrow{\beta}_y \qquad (11)$$

Backward Output Predictions

$$\overleftarrow{y_t} = W_{\overleftarrow{h}y} \cdot \overleftarrow{h}_t + \overleftarrow{\beta}_y \qquad (12)$$

Attention Layer



Queries, Keys, Values of the Attention Layer

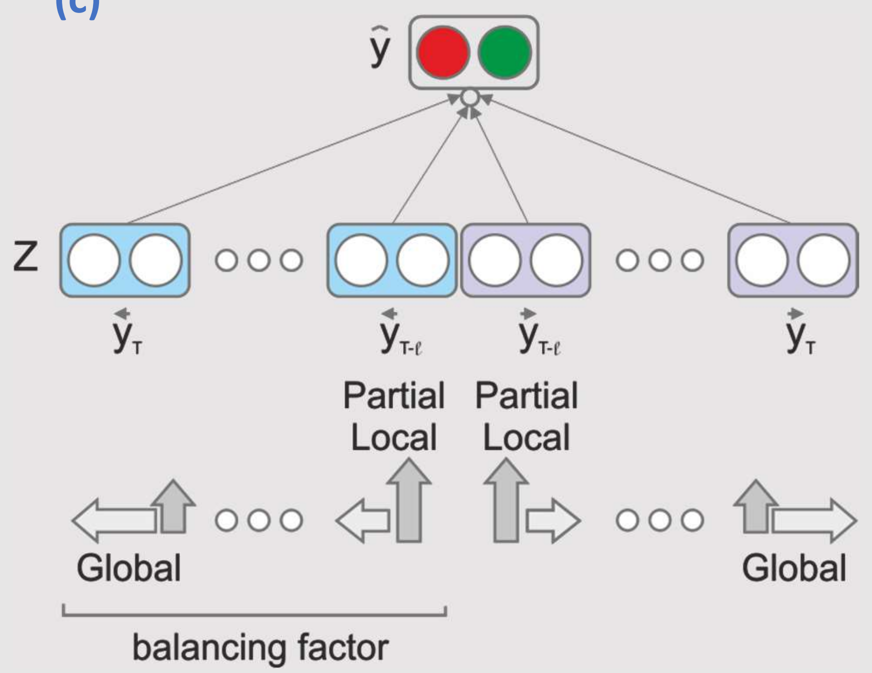
$$Q = y_t \times W_q,\ \ K = y_t \times W_k,\ \ V = y_t \times W_v \qquad (13)$$

Attention Layer

$$Att(y_t) = softmax\left(\frac{Q \times K^T}{\sqrt{h}}\right) \times V \qquad (14)$$

Classical Layer & Prediction

(c)



Classical Layer

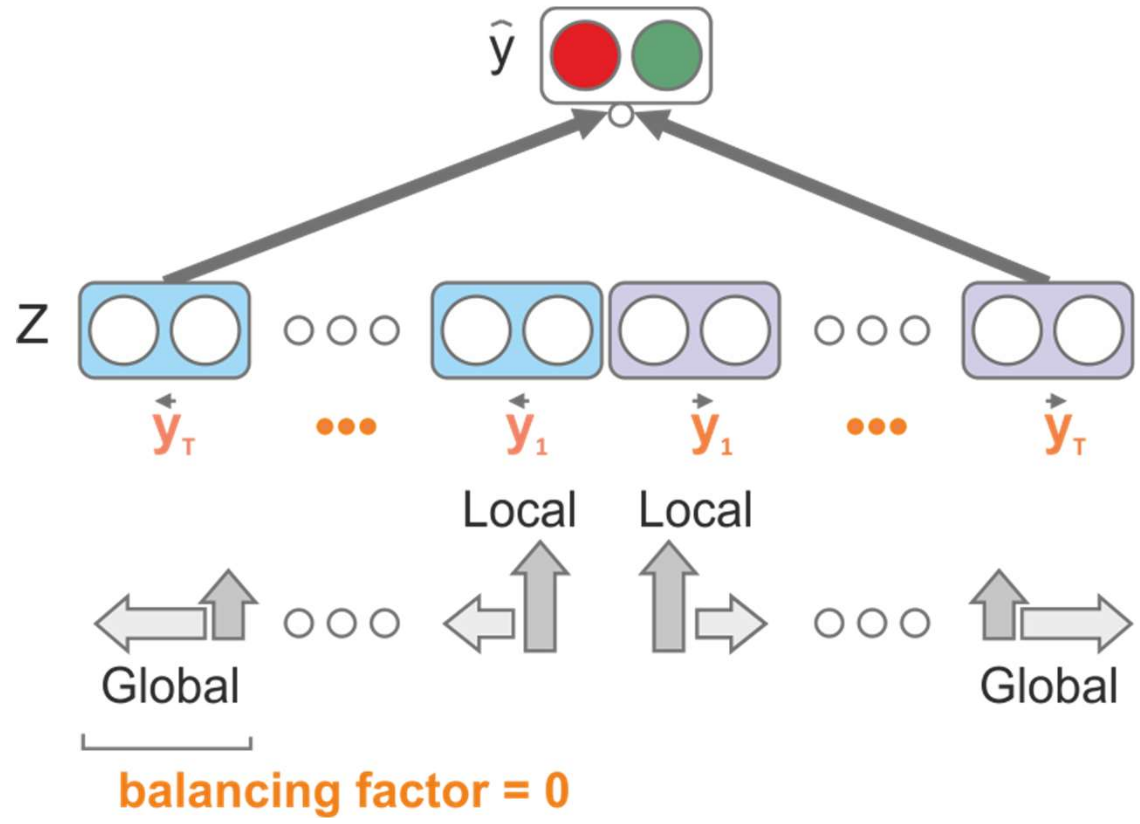
$$Z = [\overleftarrow{y}_T; \cdots; \overleftarrow{y}_{T-l}; \overrightarrow{y}_{T-l}; \cdots; \overrightarrow{y}_T] \qquad (15)$$

Prediction Layer

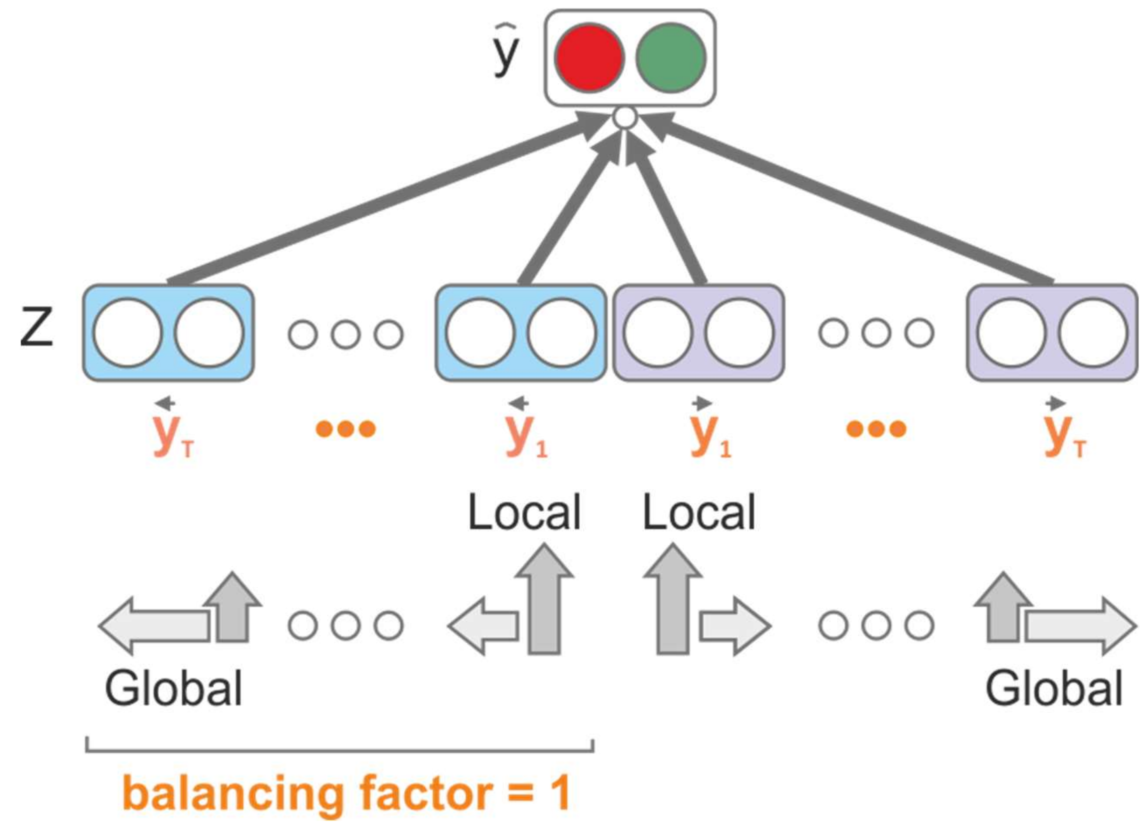
$$\hat{y} = Z \cdot w_w^T + \beta_w \qquad (16)$$

Balancing Factor and Sentiment Assigning Strategy

Cumulative Strategy



Holistic Strategy



HolC can better grasp sentiment Fluctuations

Attention, improved the generalization in all benchmark datasets

Sentiment Classification

	SUBJ	YELP-bin	SEMEVAL	SST-bin	MR	YELP	SST
HolCmax	94.80%	**91.16%**	84.96%	86.27%	84.01%	**52.75%**	49.28%
HolC	94.11%	**90.00%**	84.64%	85.48%	79.78%	**50.75%**	48.62%
HolC w/o Att	93.54%	89.54%	84.10%	84.51%	79.69%	49.46%	47.52%
CNN-static (*)	**50.51%**	**77.82%**	**71.28%**	**83.51%**	**78.30%**	**31.19%**	**43.28%**
BLSTM (*)	49.87%	75.57%	70.19%	84.34%	77.48%	31.03%	43.52%
BiLSTM-Max	92.40%	-	-	84.60%	81.10%	-	-
DAN	-	-	-	86.30%	-	-	47.70%
DCNN	-	-	-	86.80%	-	-	48.50%
CNN	93.40%	-	-	87.20%	81.50%	-	48.00%
RecNTN	-	-	-	85.40%	-	-	45.70%
CT-LSTM	-	-	-	88.00%	-	-	51.00%
C-LSTM	-	-	-	87.80%	-	-	49.20%
SWEM-concat	93.00%	-	-	84.30%	78.20%	-	46.10%
RNN-Capsule	-	-	-	-	83.80%	-	49.30%
MEAN	-	-	-	-	**84.50%**	-	**51.40%**
AdaSent	**95.50%**	-	-	-	83.10%	-	-
USE	93.90%	-	-	87.21%	81.59%	-	-
Fast Dropout	93.60%	-	-	-	-	-	-
SDAE	90.80%	-	-	-	74.60%	-	-
GRU-RNN	91.85%	-	-	-	78.26%	-	45.02%
Capsule-B	93.80%	-	-	86.80%	82.30%	-	-
Emo2Vec	-	-	-	82.30%	-	-	43.60%
BiLSTM-CRF & CNN	-	-	-	**88.30%**	82.30%	-	48.50%
SwissCheese	-	-	82.00%	-	-	-	-
CUFE	-	-	83.40%	-	-	-	-
ECNU	-	-	84.30%	-	-	-	-
UNIMELB	-	-	**87.00%**	-	-	-	-
Thecerealkiller	-	-	82.30%	-	-	-	-
TwiSE	-	-	82.60%	-	-	-	-
Finki	-	-	84.80%	-	-	-	-

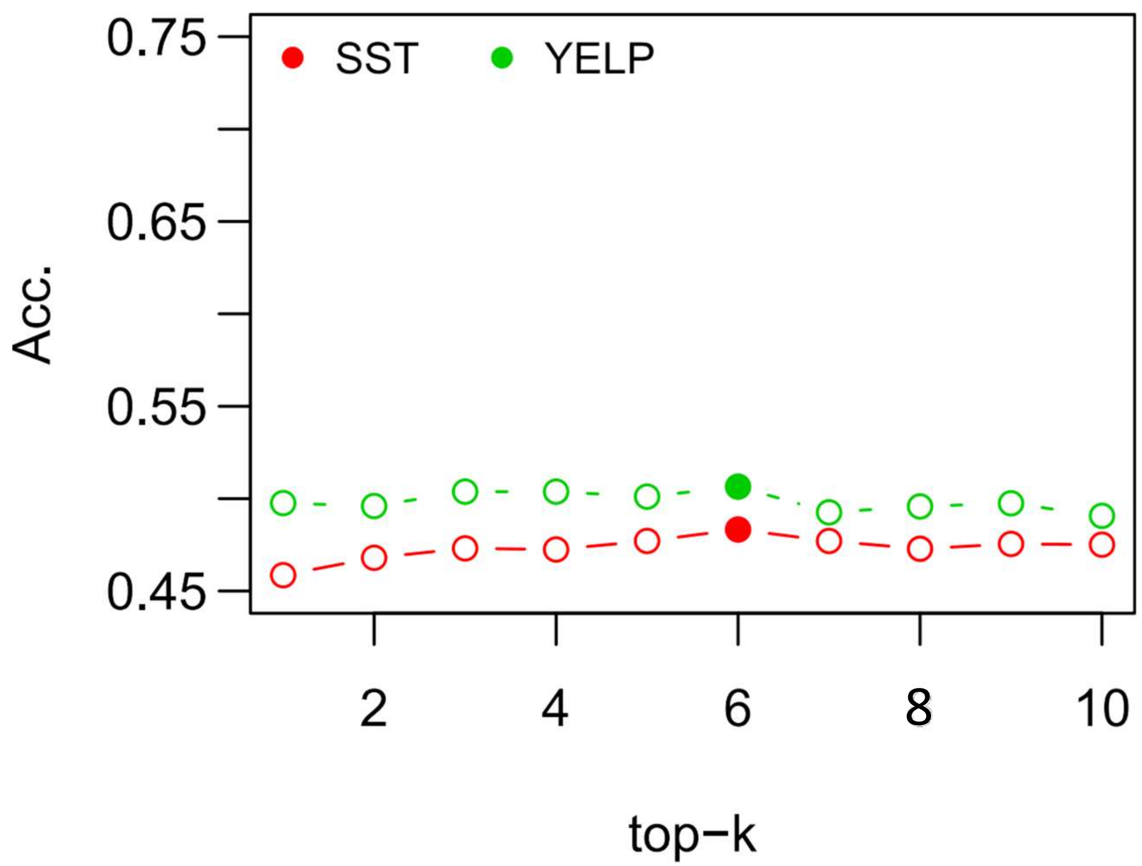
Question Type Classification

TREC dataset

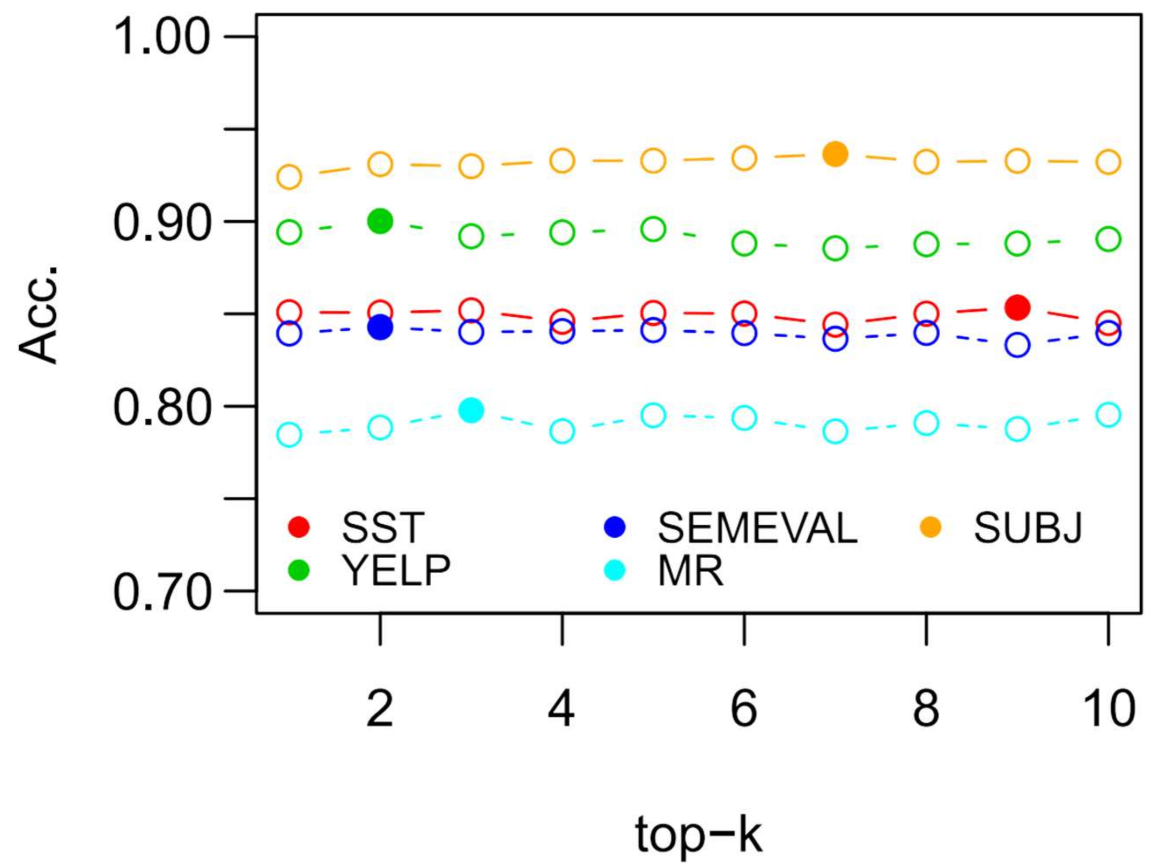
	Acc	Best Acc
HolC	98.64%	**99.00%**
HolC w/o Att	98.48%	**98.80%**
CNN-static (*)	98.60%	98.80%
BLSTM (*)	97.84%	98.80%
CNN	93.60%	-
AdaSent	92.40%	-
BiLSTM-Max	88.20%	-
DCNN	93.00%	-
USE	98.07%	-
SDAE	78.40%	-
GRU-RNN	93.00%	-
Capsule-B	92.80%	-
SWEM-aver	92.20%	-

The Contribution of the k-max Pooling Operation

HolC (with Attention) Accuracy in fine- grained datasets for various values of k

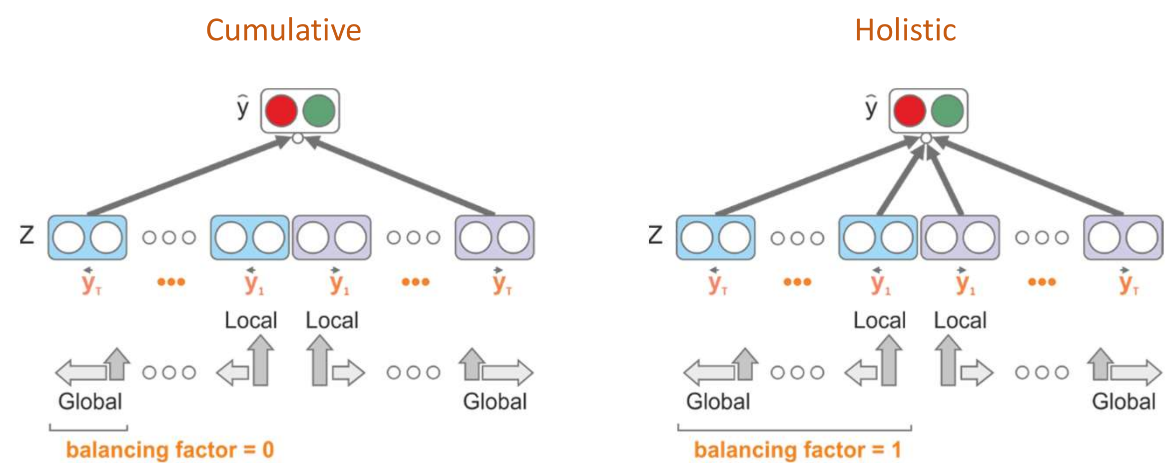


HolC (with Attention) Accuracy in Binary datasets for various values of k.



BF	SUBJ	YELP-bin	SEMEVAL	SST-bin	MR	YELP	SST	TREC
0	**94.11%**	89.67%	**84.64%**	**85.48%**	**79.78%**	**50.75%**	**48.62%**	98.60%
0.25	93.38%	89.66%	84.09%	84.95%	79.47%	50.19%	47.13%	98.44%
0.5	92.23%	89.80%	84.28%	84.63%	79.03%	49.16%	47.65%	98.52%
0.75	93.28%	89.52%	84.19%	85.10%	78.73%	50.65%	47.78%	**98.64%**
1.00	93.50%	**90.00%**	84.31%	84.96%	79.22%	49.85%	47.04%	98.60%

Holistic/Cumulative content identification over a set of different datasets & balancing factor (BF)



Conclusion - Novelties of the Proposed Method

(1) The introduction of a sentence embedding via a Convolution Neural Network

(2) A bi-directional recurrent neural network for encoding semantic content sequentially

(3) Classical layer that exploits both local and global information

(4) A hyperparameter that balances mixed content motifs named Balancing Factor

(5) An improved convolution operation that better exploits the input information

(6) A k-max-pooling operation over the single max-pooling after the convolution layer

(7) An improved design of the attention layer capable of improving the generalization task and

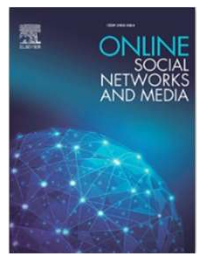
(8) The utilization of pre-trained word vectors over the randomly initialized ones

More…

Thank you for your attention!

Pantelis Agathangelou, Ioannis Katakis

Balancing Between Holistic and Cumulative Sentiment Classification



HolC – source code: github.com/unic-ailab/Holistic-Cumulative