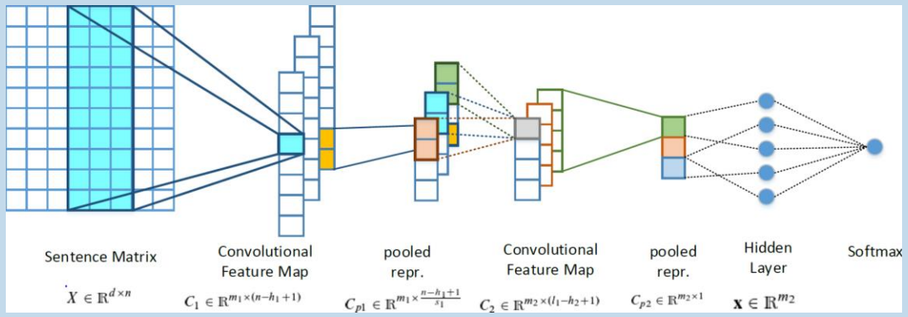
SwissCheese: Sentiment Classification using an Ensemble of Convolutional Neural Networks and Distant Supervision

Jan Deriu[1], Maurice Gonzenbach[1], Fatih Uzdilli[3], Aurelien Lucchi [1,] Valeria De Luca [2,] Martin Jaggi [1]
[1] Data Analytics Laboratory , ETH Zurich; [2] Computer Vision Laboratory , ETH Zurich; [3] Institut für angewandte Informationstechnologie , ZHAW

CNN 2Layer Architecture



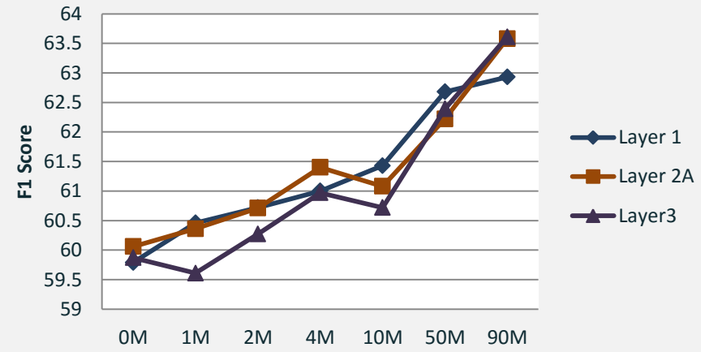
3 Step Training

Word Embeddings
- S1: word2vec with 200M tweets
- S2: GloVe with 90M tweets
- Time: 3-4 days

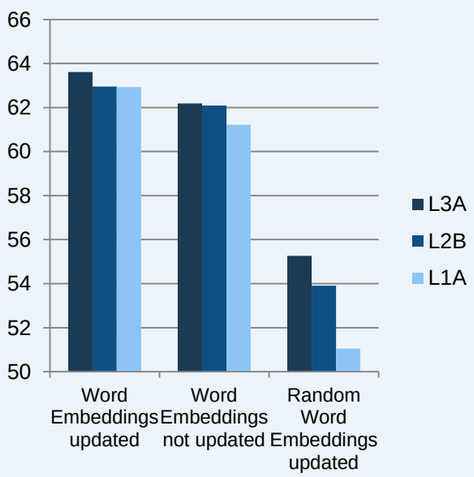
Distant Phase
- S1: 90M tweets
- S2: 60M tweets
- Time: 10-15 hours

Supervised Phase
- Both: 18K provided training data
- Time: up to 1h

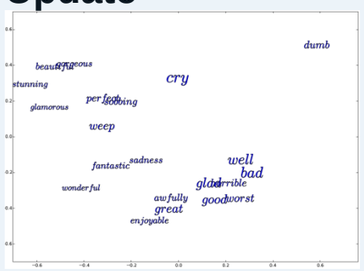
Number of Tweets in the Distant Phase



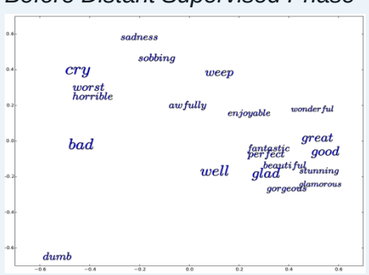
Importance of high-quality Word Embeddings



Word Embeddings Update



Before Distant Supervised Phase



After Distant Supervised Phase

Meta Classifier

Supervised Phase: High Variance in F1-score over # epochs
Goal: Increase Robustness
Solution: Train a Random Forest on the outputs of the various systems

S1: Trained for different number of epochs (a-f)
S2: Trained until it reached good average scores among validation sets

Technical Details

Number of Kernels: *Both*: 200
Filter Lengths: *S1*: h_1=6, h_2=3 *S2:* h_1=6, h_2=4
Pooling Length *S1*: w_1=6, st_2=2 *S2:* w_1=3, st_2=3

Optimization: AdaDelta, S1: no regularization S2: L2 regularization

Results

	Test 2016	Test 2015	Test 2014	Test 2013
S1a	60.47	64.26	73.98	**71.52**
S1b	62.73	65.80	74.60	70.10
S1c	61.89	64.80	**75.70**	70.90
S1d	60.58	64.20	74.15	71.50
S1e	57.19	61.02	69.12	67.00
S1f	62.20	66.70	72.00	68.00
S2	62.36	66.63	72.45	70.05
FS	**63.30**	**67.05**	71.55	70.01