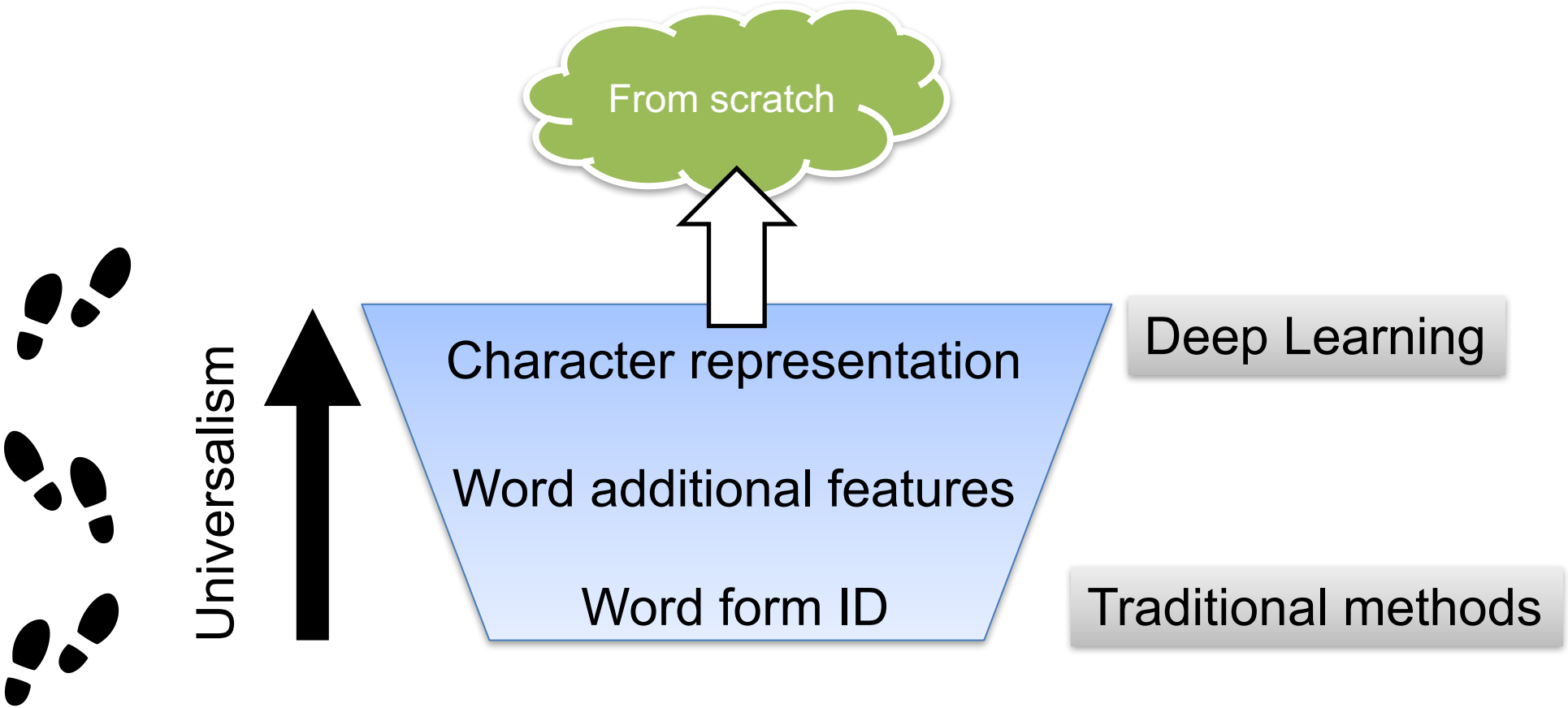
National Research Center
“Kurchatov Institute”

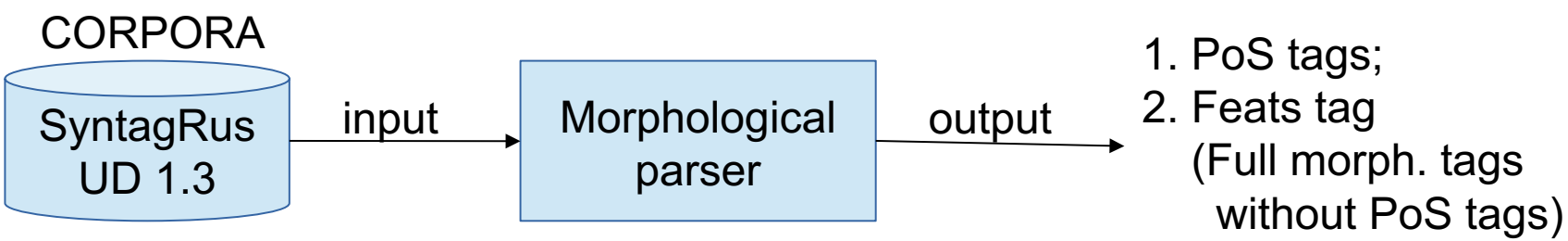
“Research of a deep learning neural network effectiveness for a morphological parser of Russian language”

Sboev A.G.

and
Gudovskikh D.V., Moloshnikov I.A., Rybka R.B.,
Voronina I. (Voronezh State University),
Ivanov I. (Moscow Technological University)

ML approaches for morphological parsing

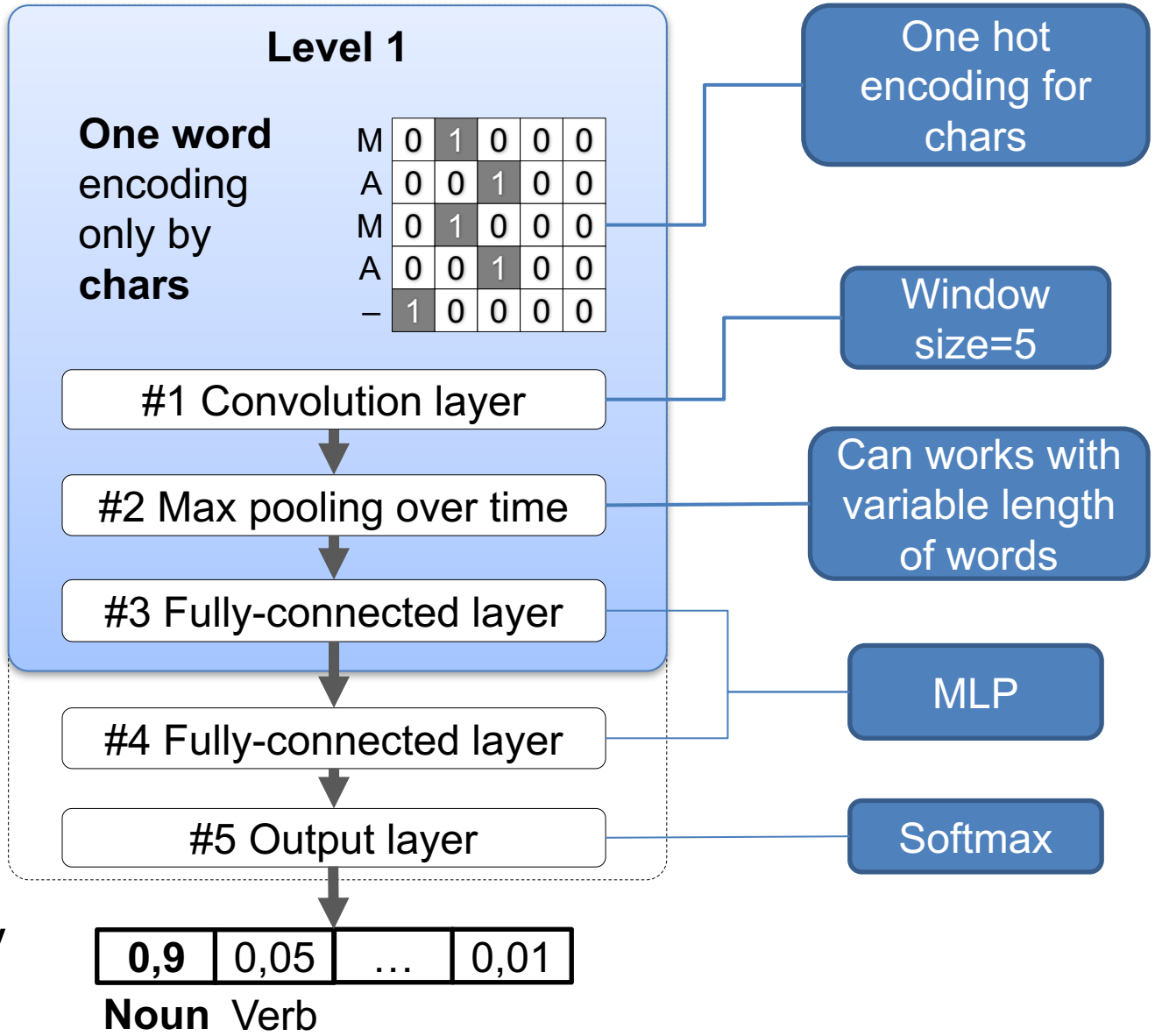




Two-level neural model, PoS tag
Level 1 pre-train

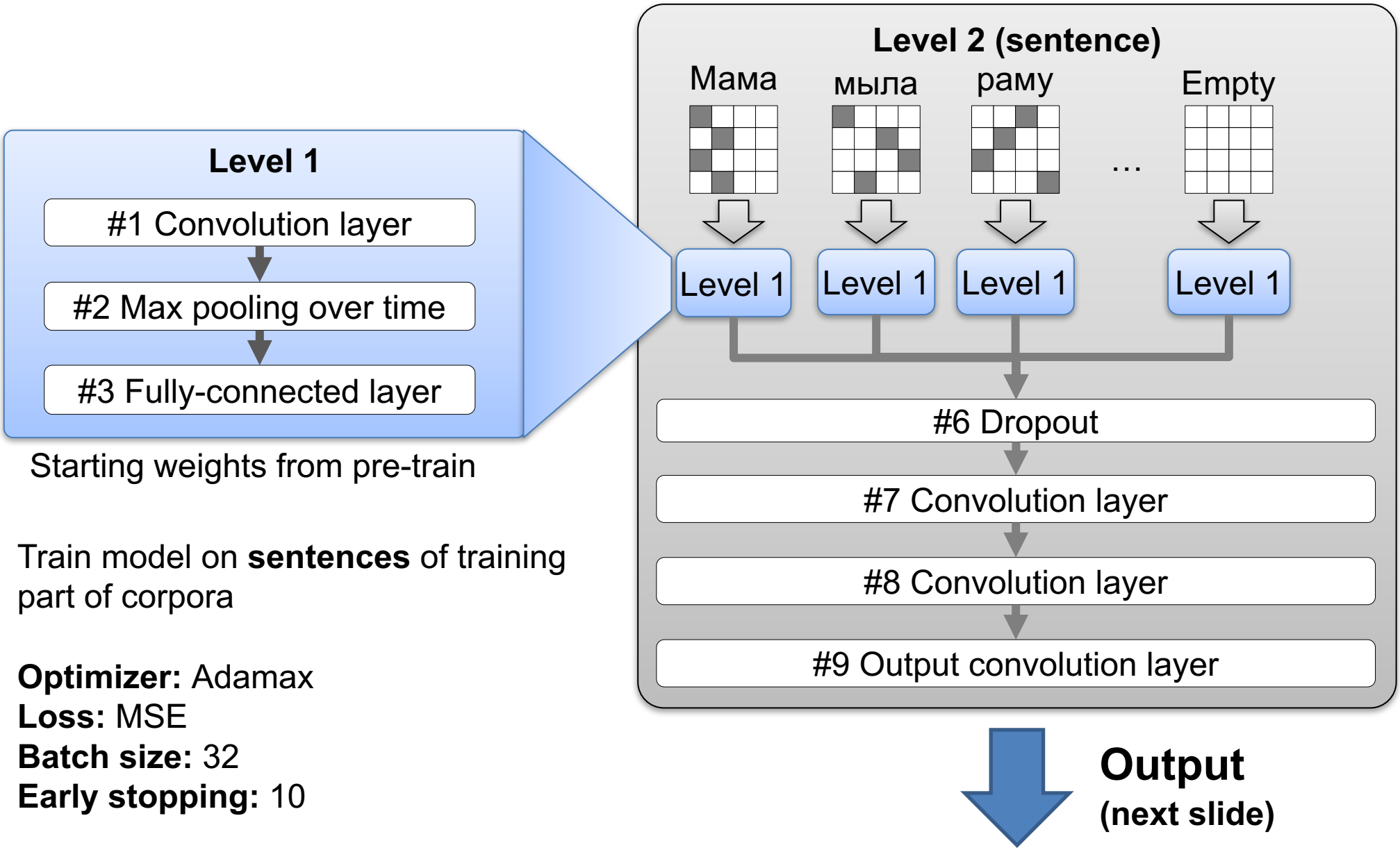
Pre train model on **each word** of training part of corpora

Optimizer: Adamax
Loss: MSE
Batch size: 1024
Early stopping: 15



PoS tags probability for **one** word

Two-level neural model, PoS tag
Level 2 training

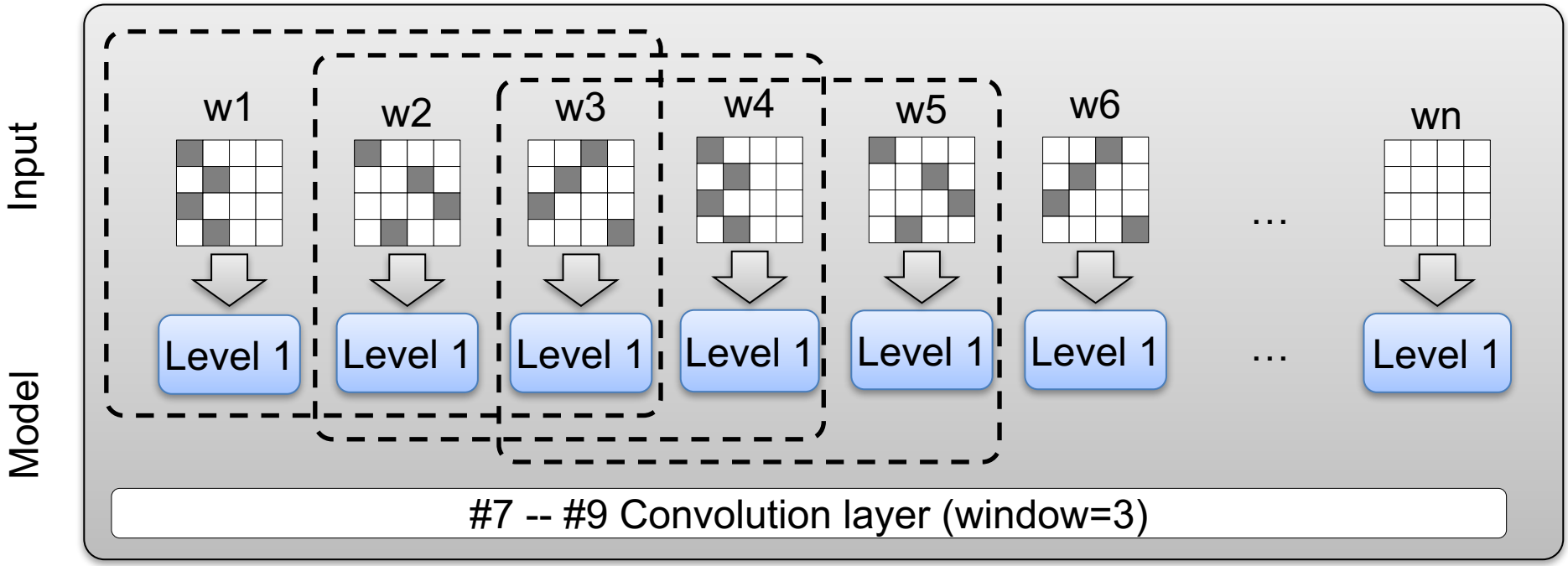


Starting weights from pre-train

Train model on **sentences** of training part of corpora

Optimizer: Adamax
Loss: MSE
Batch size: 32
Early stopping: 10

Two-level neural model, PoS tag Level 2 output

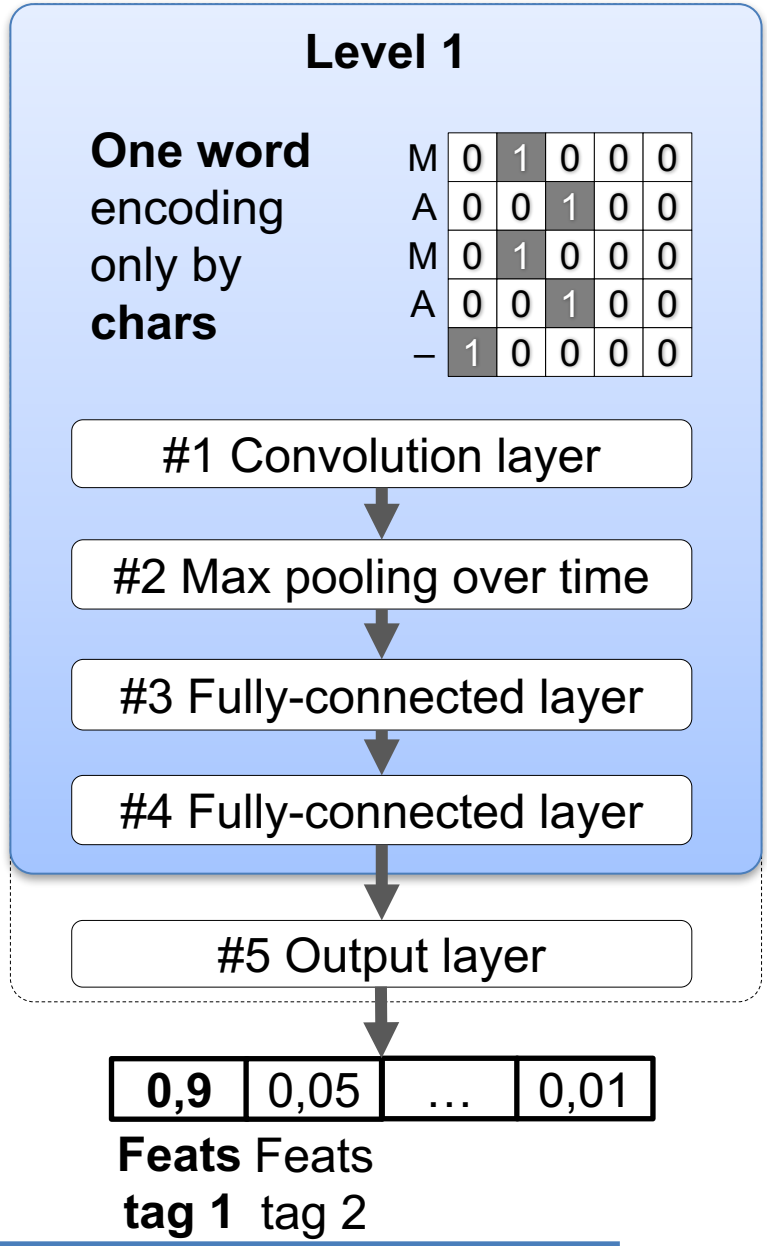


Output

Output

PoS tags	Probability of PoS tags for words of the sentence				
	w1	w2	w3	…	wn
NOUN	**0.8 (noun)**	0.02	**0.99 (noun)**		0.1
VERB	0.01	**0.7 (verb)**	0.001		0.001
ADJ	0.01	0.1	0.001		**0.51 (adj)**
…					
∑	**1**	**1**	**1**	**…**	**1**

Two-level neural model, feats tag
Level 1 pre-train



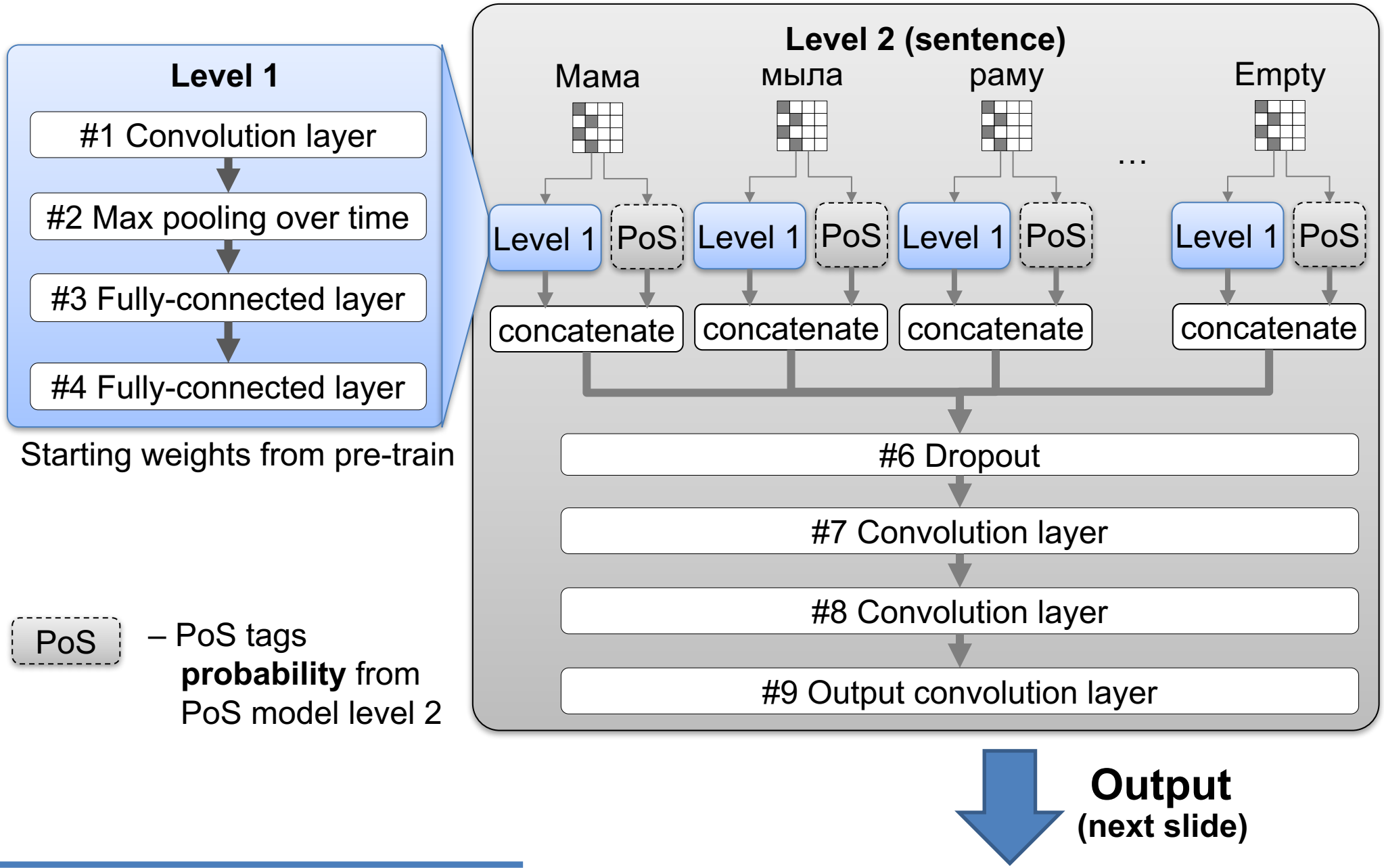
Pre train model on **each word** of training part of corpora

Optimizer: Adamax
Loss: MSE
Batch size: 1024
Early stopping: 15

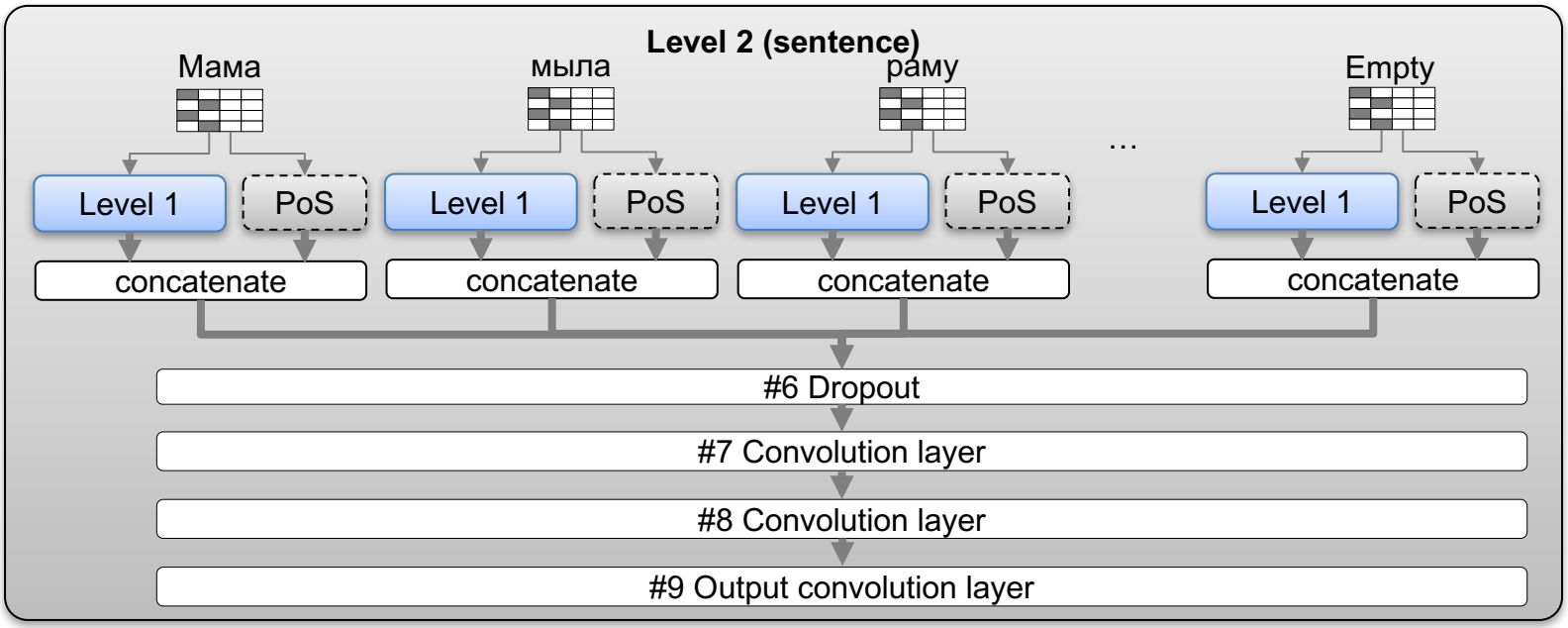
Feats tag — unique combination of morphological tags in training corpora, except PoS.

Example:
{Animacy=Inan|Case=Nom|Gender=Masc|Number=Sing}

Two-level neural model, feats tag
Level 2 training



Two-level neural model, feats tag Level 2 output



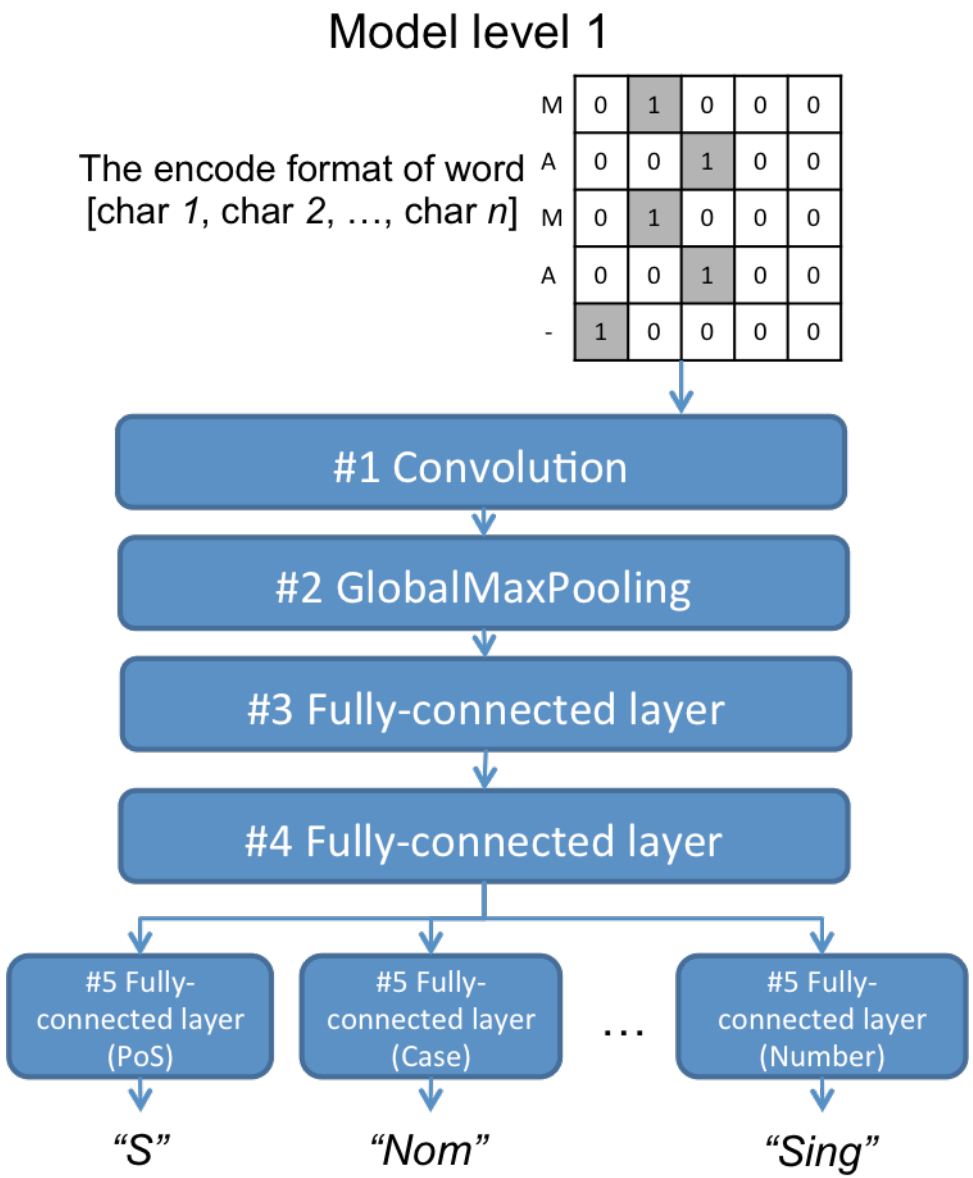
Output

Output

Feats tags	Probability of feats tags for words of the sentence				
	w1	w2	w3	…	wn
Feats tag 1	**0.8 (tag 1)**	0.02	**0.99 (tag 1)**		0.1
Feats tag 2	0.01	**0.7 (tag 2)**	0.001		0.001
Feats tag 3	0.01	0.1	0.001		**0.51 (tag 3)**
…					
∑	**1**	**1**	**1**	**...**	**1**

One-level deep learning model

Predict morphology tags based on **one word** as sequence of chars



Models for comparison Traditional ML approaches

- SVM-based Approach;
  - N-gram with one-hot-encoded words
  - the ensemble of linear SVM base of one-vs-all strategy

- Extension of the SVM approach;
  - Add morphology **hypothesis** from Mystem:
    - Possible full tags for each word from Mystem;
    - Possible morphological features for each unparsed word from n-gram
  - Add full tag **scoring function**

Evaluation

- **SynTagRus** corpora in:
  - original format of National Russian Corpus
  - Universal dependencies 1.3 format
- Metrics
  - Accuracy
  - F1-score weighted

datasets type	Number of PoS features	Number of morphological features
Original	11	45
UD-1.3	15	36
UD-1.4	16	36

Results SynTagRus original format

Model name	all		Out-of-vocabulary	
	PoS	Full	PoS	Full
	accuracy	accuracy	accuracy	accuracy
LinearSVC: window size (3,5,7)	94.10 - 95.02	83.9 - 85.74	32.11 - 63.33	29.7 - 30.2
LinearSVC+Mystem (win=8)	95.61	81.65	95.91	79.6
One-level model	96.63	85.58	94.72	74.76
(Proposed approach)	98.24	94.12	95.14	84.4
(Proposed approach + Dropout)	98.34	94.83	95.24	85.07

Results SynTagRus
Universal Dependencies 1.3 format

Model name	all		Out-of-vocabulary	
	PoS	Full	PoS	Full
	accuracy	accuracy	accuracy	accuracy
LinearSVC: window size (3,5,7)	94.87-95.46	82.3-84.04	68.85-69.22	11.91-13.32
One-level model	96.85	85.56	94.13	59.86
(Proposed approach)	98.44	93.34	95.16	71.3
(Proposed approach + Dropout)	98.49	94.31	95.07	74.48
Google SyntaxNet	98.27	94.01	94.21	74.12

MorphoRuEval on Dialog 2017 Closed track

Place	team ID	Accuracy by tags	Accuracy by sentences
1	C	93,39	65,29
2	O	93,08	62,71
3	H (sagteam)	92,64	58,40

Conclusion

- **Character representation** along with deep learning models have high potential

- The Mystem **hypotheses** is effective in case of SVM models in case of out-of-vocabulary words.

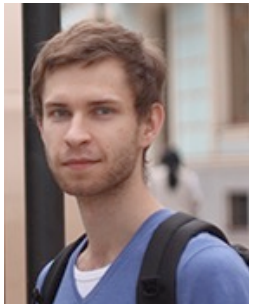
- For practical needs it would be useful **to unite** these approaches in common morphological parser increase the universality of parsing with a higher accuracy.

Acknowledgments
The study was funded by RFBR according to the research project № 16-37-00214

Thank you for your attention!
Any questions?

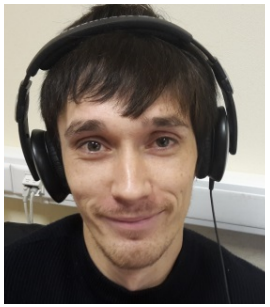
Sboev Alexander
sag111@mail.ru

A.Sboev



R.Rybka

I.Moloshnikov



D.Gudovskikh

I.Ivanov

I. Voronina