



NATIONAL RESEARCH
UNIVERSITY

Mitrofanova Marina

Отвeты MAIL.RU

Kaggle соревнование

Moscow, 2018

TASK - Classification

Metrics for evaluation - **f1-score**

Predict main category of the question

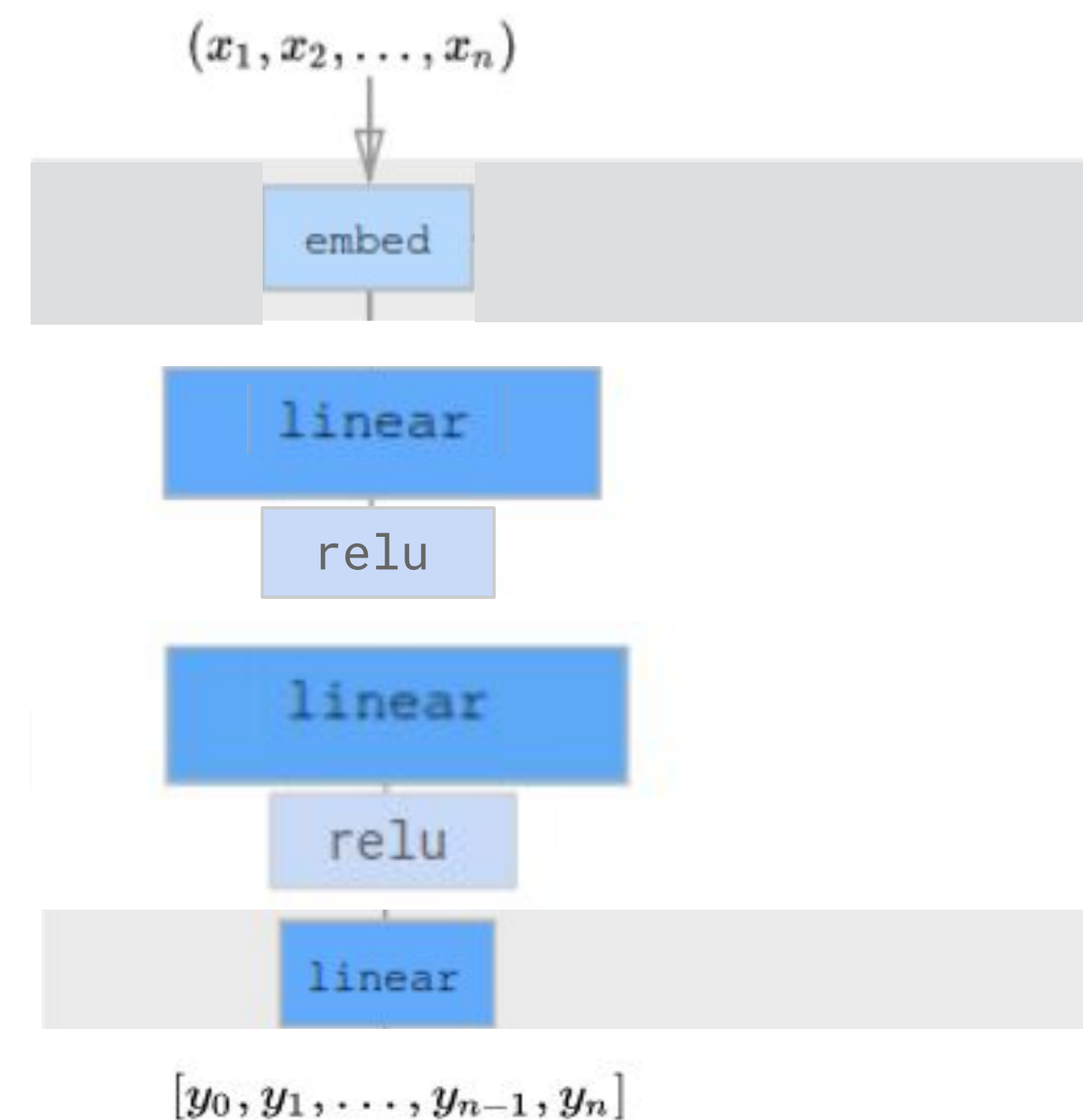
main_category	sub_category	question
4	32	Законно ли изменение количества дней листа нет...
1	8	Погадайте что будут, у меня предчувствия?) К ч...
10	66	Какой дровокол надёжней?

BASELINE

Model

Thematic title of the main text

1. FastText Embeddings
2. Two Linear layers



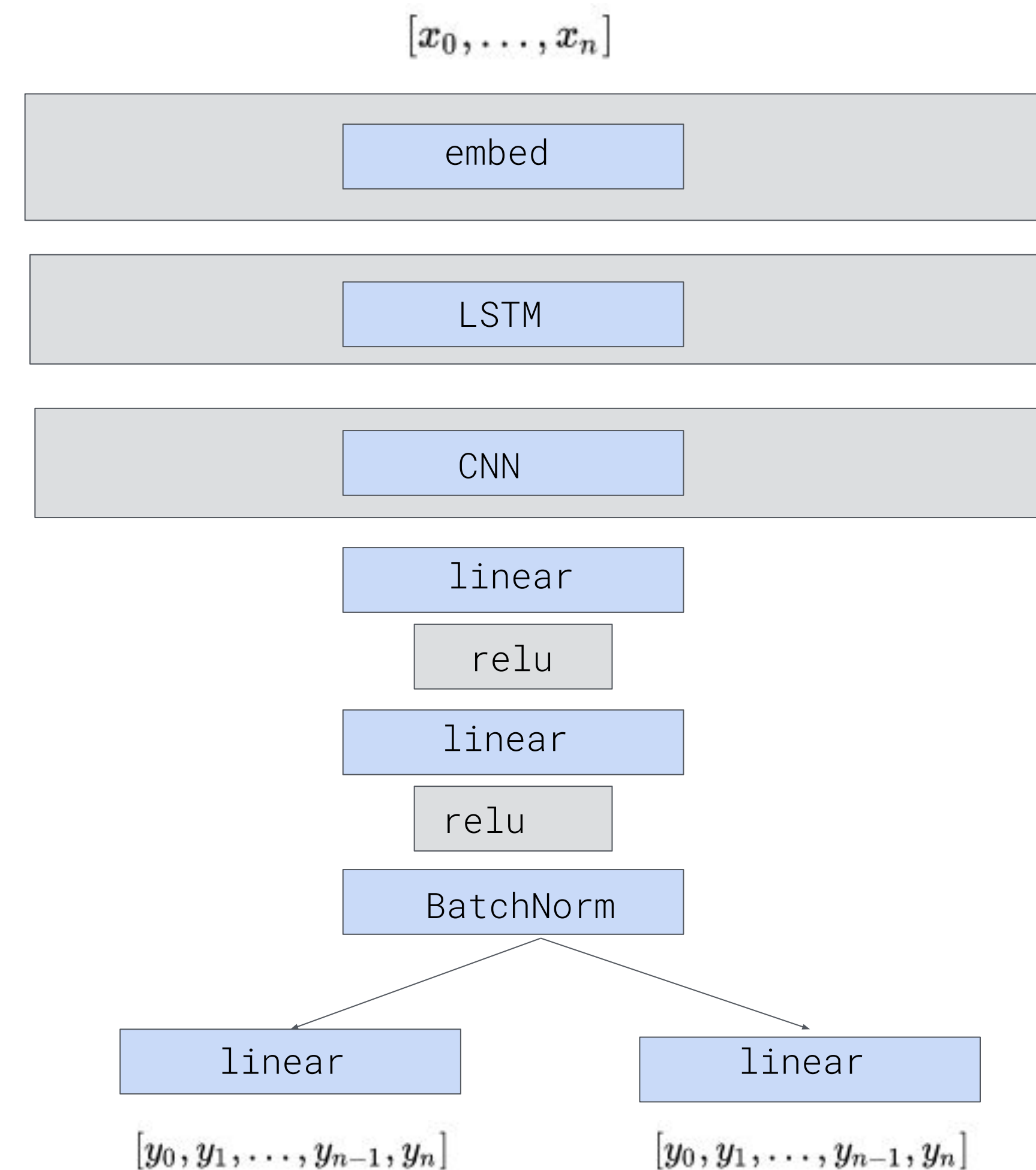
SOLUTION

Presentation subtitle

Multitask

Learning

- + Weighted training
- + Dropout
- + BatchNorm
- + Additional layers:
 - LSTM
 - CNN
 - BatchNorm



$$\text{loss} = 0.6 * \text{loss1} + 0.4 * \text{loss2}$$

RESULTS

- | model | f1-score |
|---------------------|-------------------------|
| • Baseline | - 0.56137 |
| • Our Model | - <u>0.61843</u> |
| • Best Model | - 0.64554 |



NATIONAL RESEARCH
UNIVERSITY

Questions?