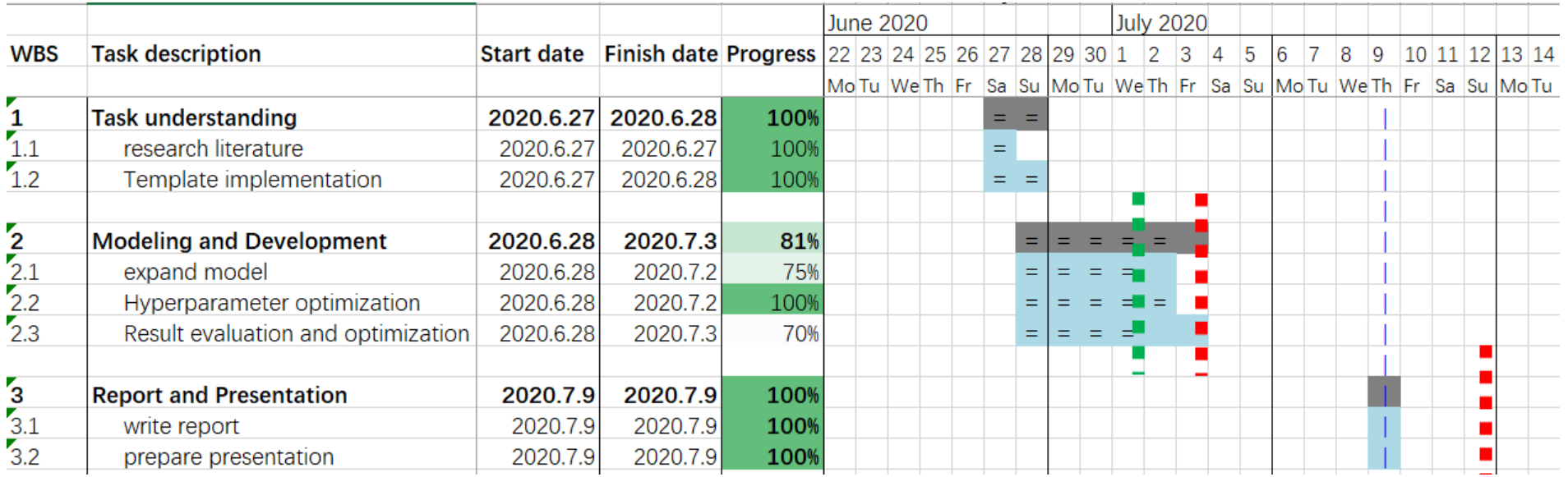
Deep Learning for NLP 2020

Shared Task

Group 30, Yi Cui, Zixuan Chen

WBS	Task description	Start date	Finish date	Progress
1	**Task understanding**	**2020.6.27**	**2020.6.28**	**100%**
1.1	research literature	2020.6.27	2020.6.27	100%
1.2	Template implementation	2020.6.27	2020.6.28	100%
2	**Modeling and Development**	**2020.6.28**	**2020.7.3**	**81%**
2.1	expand model	2020.6.28	2020.7.2	75%
2.2	Hyperparameter optimization	2020.6.28	2020.7.2	100%
2.3	Result evaluation and optimization	2020.6.28	2020.7.3	70%
3	**Report and Presentation**	**2020.7.9**	**2020.7.9**	**100%**
3.1	write report	2020.7.9	2020.7.9	**100%**
3.2	prepare presentation	2020.7.9	2020.7.9	**100%**



43	qwerqwer	5	07/03/20	0.2960 (41)
44	m-bodensohn	1	07/01/20	0.2925 (42)

Agenda

TECHNISCHE UNIVERSITÄT DARMSTADT

- Task
- Model
- Result

NLLG

Task

- Semantic similarity:
  - a metric defined over a set of documents or terms, where the idea of distance between items is based on the likeness of their meaning or semantic content as opposed to lexicographical similarity [1]
- Adversarial attacks:
  - **disemvoweling**: Some of the vowels have been removed. [2]
  - **visual attacks**: Similar-looking characters have been replaced by each other. [2]

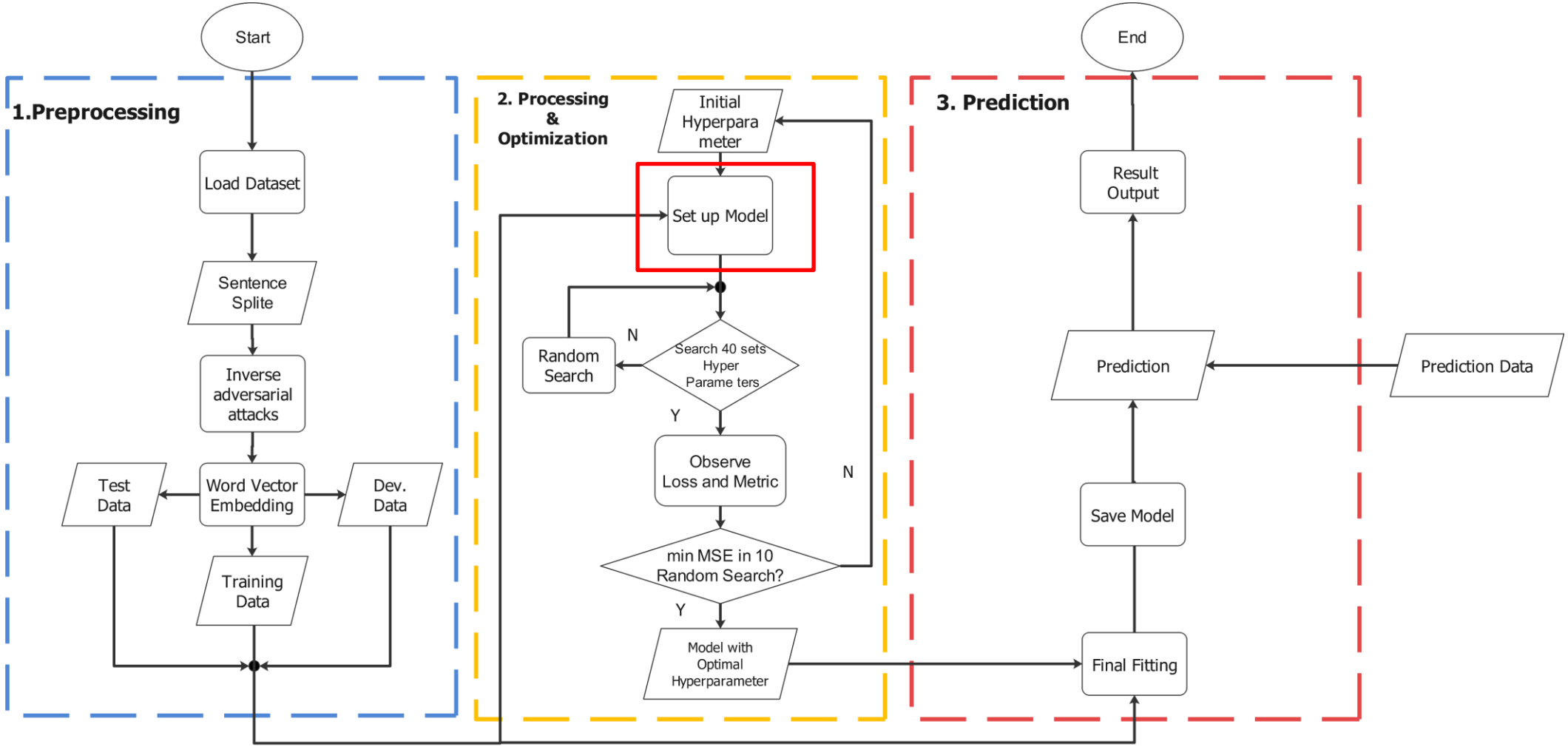
Three dgs playing in th snow. ⟷ 0.84 ⟷ ʔhreɛ ɒogs run in the snoᑕᑐ.

Task

TECHNISCHE UNIVERSITÄT DARMSTADT

- Approach
  - Preprocessing
    - **inverse adversarial attacks** → reduce OOV
    - **Word vector** → Embedding words with wiki-news-300d-1M.vec
  - Processing
    - apply MLP with **Random Search** → find appropriate mapping
    - **extend MLP with LSTM** → prognosticate semantic similarity with context

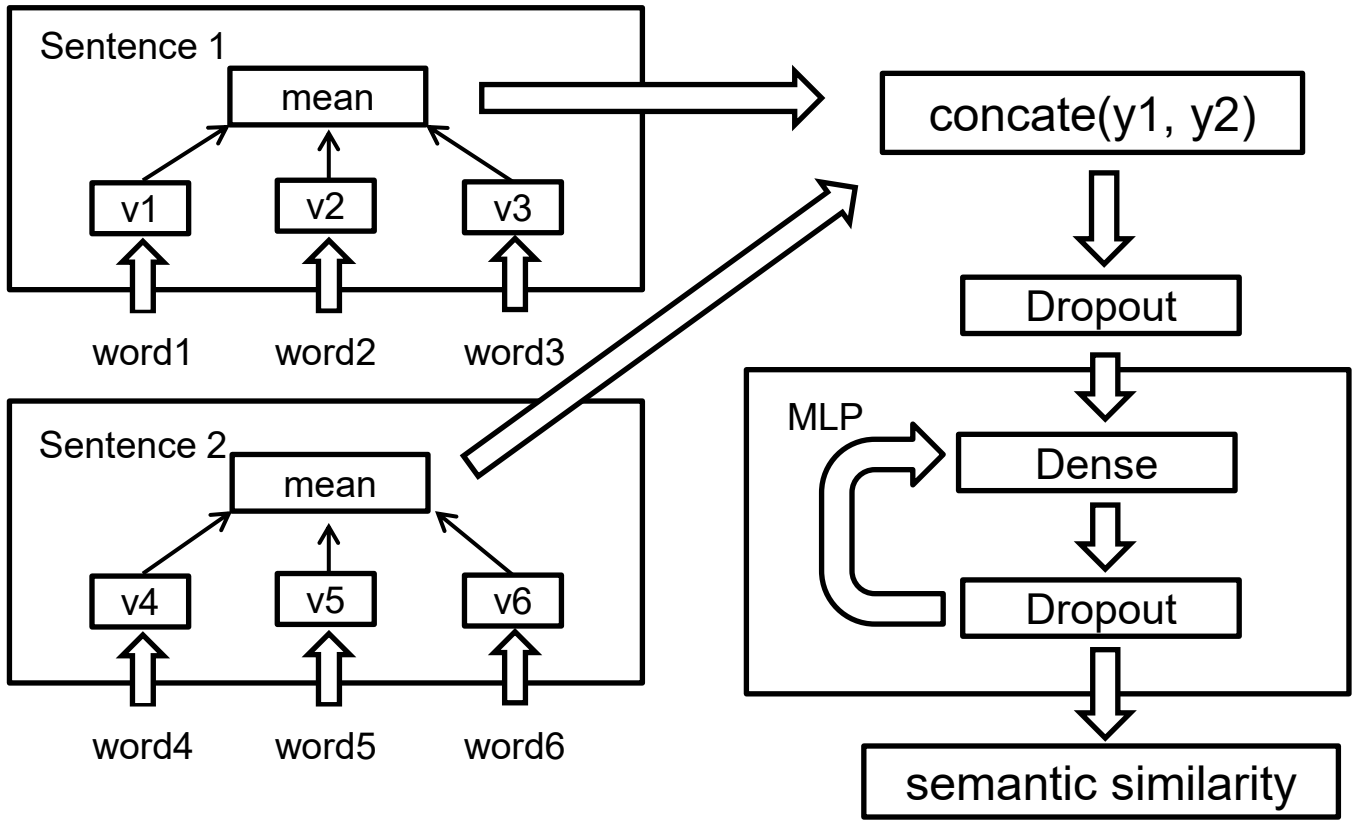
Model (Chart Flow)



Model (MLP)

- Hyper Parameter

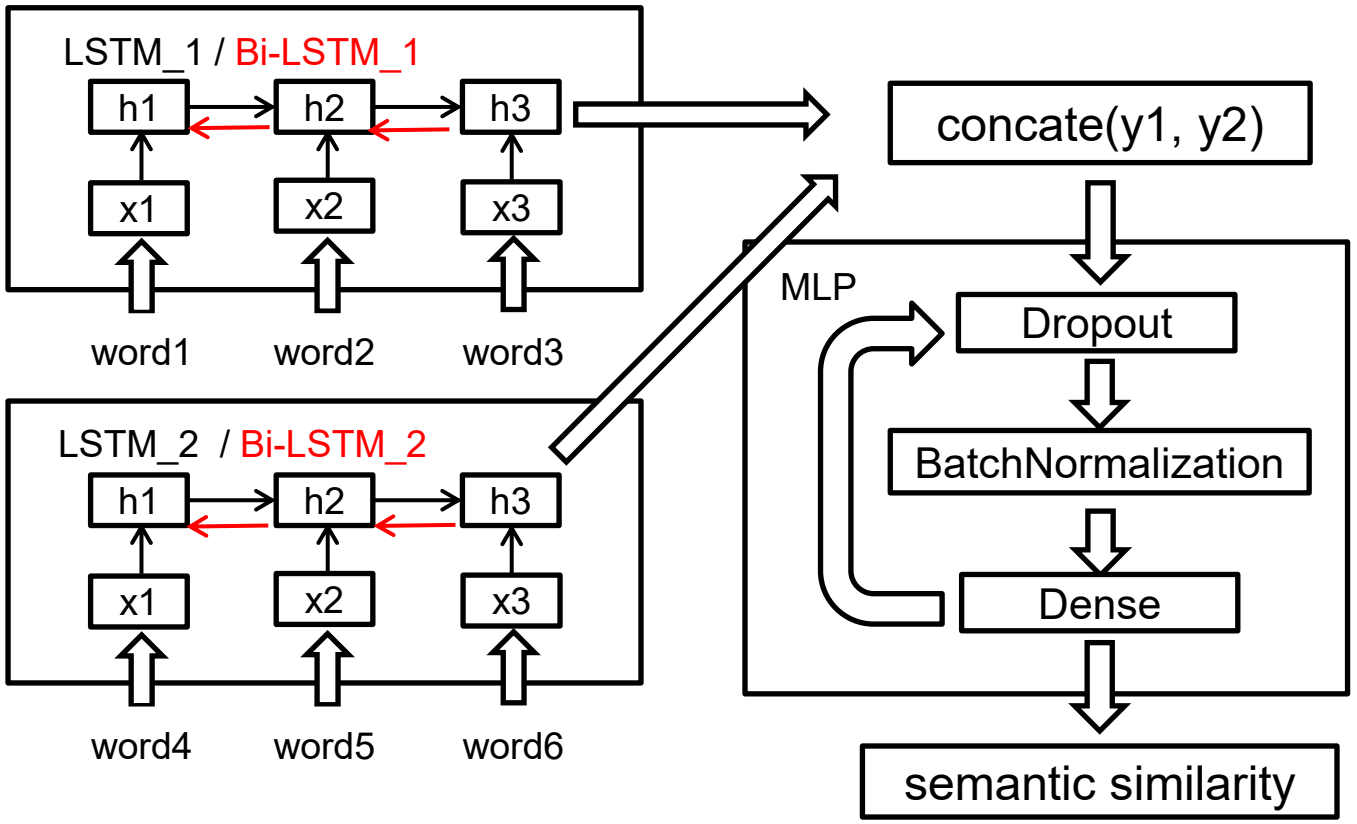
Type	Value
Dense size	337, 171
Dropout rate	0.5, 0.1, 0.2
Activation	‘relu’, ‘selu’, ‘selu’
Loss	MSE
Metric	MSE



Model (LSTM and Bi-LSTM)

- Hyper Parameter

Type	Value
Emb_size	56
LSTM Units	300
Dense size	440, 201, 107, 58,
Dropout rate	0.2, 0.2, 0.1, 0.2, 0.2
Activation	‘selu’, ‘hard sigmoid’, ‘relu’, ‘hard sigmoid’,
Loss	MSLE
Metric	MSE



Overview of result

TECHNISCHE UNIVERSITÄT DARMSTADT

- Score from CodaLab:

MLP > LSTM > Bi-LSTM

	MLP	LSTM + MLP	Bi-LSTM + MLP
Development	0.3982	0.3092	0.2844
Test	0.2960	None	None

- Hidden Layer (Dense) less than 3 → Otherwise Overfitting

- Model is sensitive to Data (OOV)

  → inverse attacks are more important than modeling

- Time management are essential

Reference

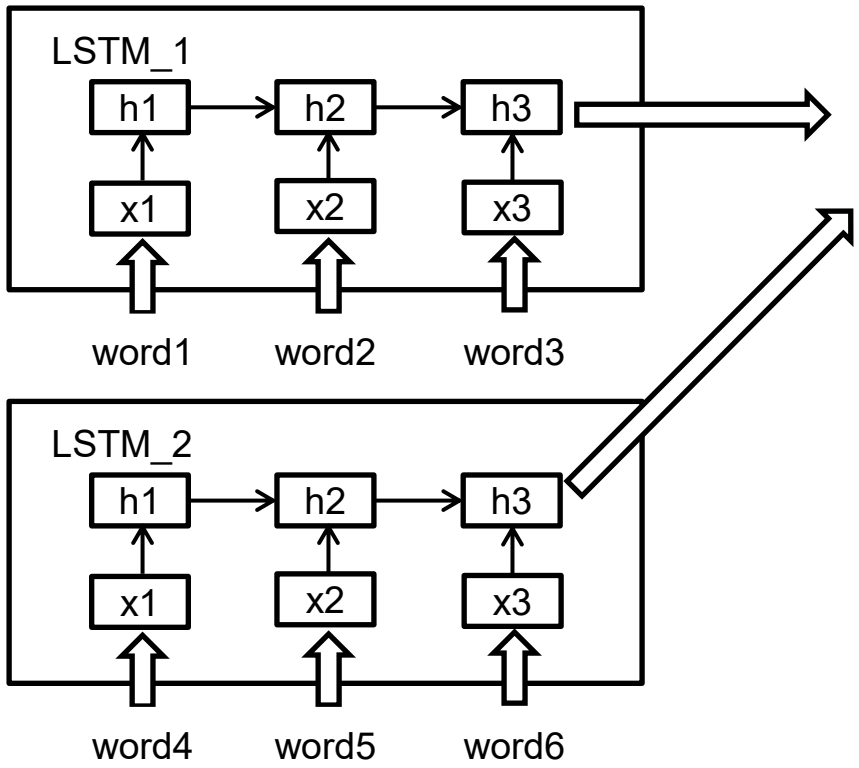
- [1]: https://en.wikipedia.org/wiki/Semantic_similarity
- [2]: Deep Learning for NLP 2020 Shared Task Organization

TECHNISCHE UNIVERSITÄT DARMSTADT

Thanks for your attention

NLLG

Back up


LSTM_1
h1
h2
h3
x1
x2
x3
word1
word2
word3
LSTM_2
h1
h2
h3
x1
x2
x3
word4
word5
word6