Literature review of SOTA Transfer Learning and Multi-task Learning applications

Jiaxi Zhao, 20.04, Guided Research Presentation

Chair of Software Engineering for Business Information Systems (sebis)
Faculty of Informatics
Technische Universität München
wwwmatthes.in.tum.de

Outline

Motivation

Transfer Learning

Multi-task Learning

Outline

Goals

·Analysing TL and MTL application cases.

·Summarizing TL and MTL suitable conditions.

·Guidance for researchers to solve data scarcity with TL&MTL by writting a 75 pages research paper.

Outline

Motivation

Goals

Concept

Model Structure

Tasks

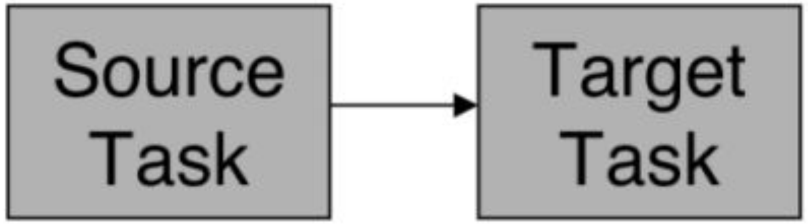
- Text
- Protein
- Source code

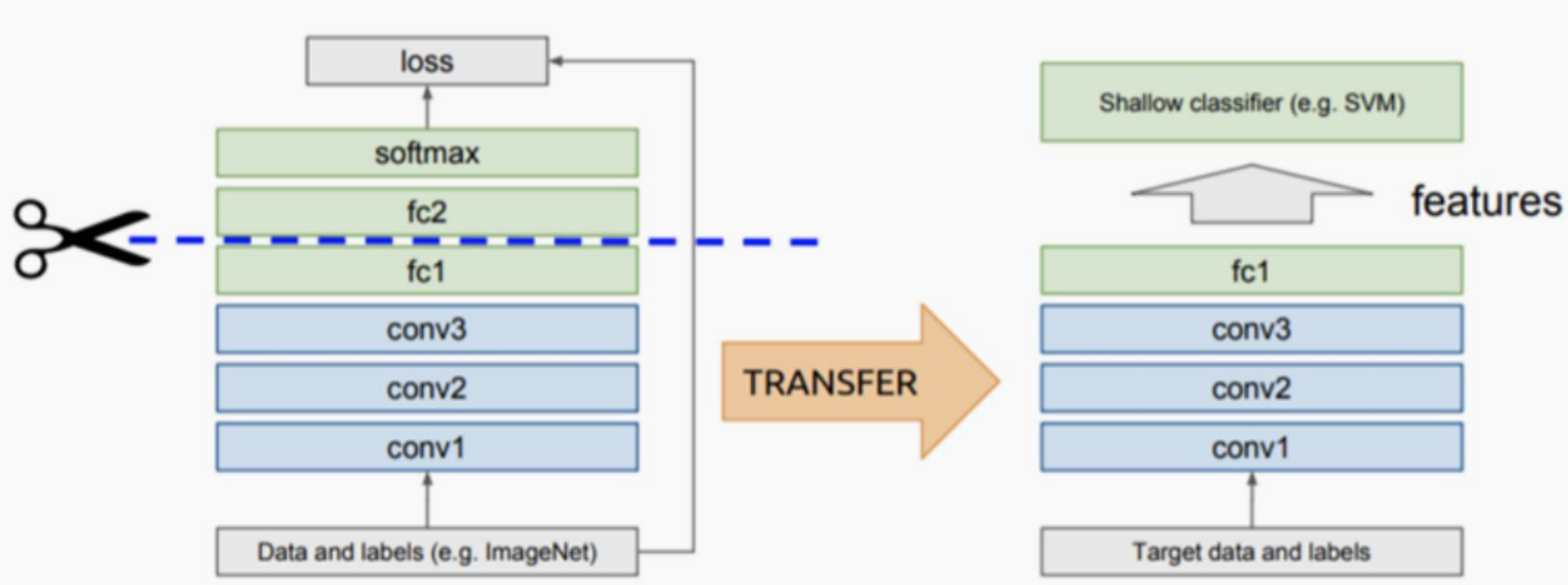
Conclusion

Outlook

Concept

Transfer Learning:

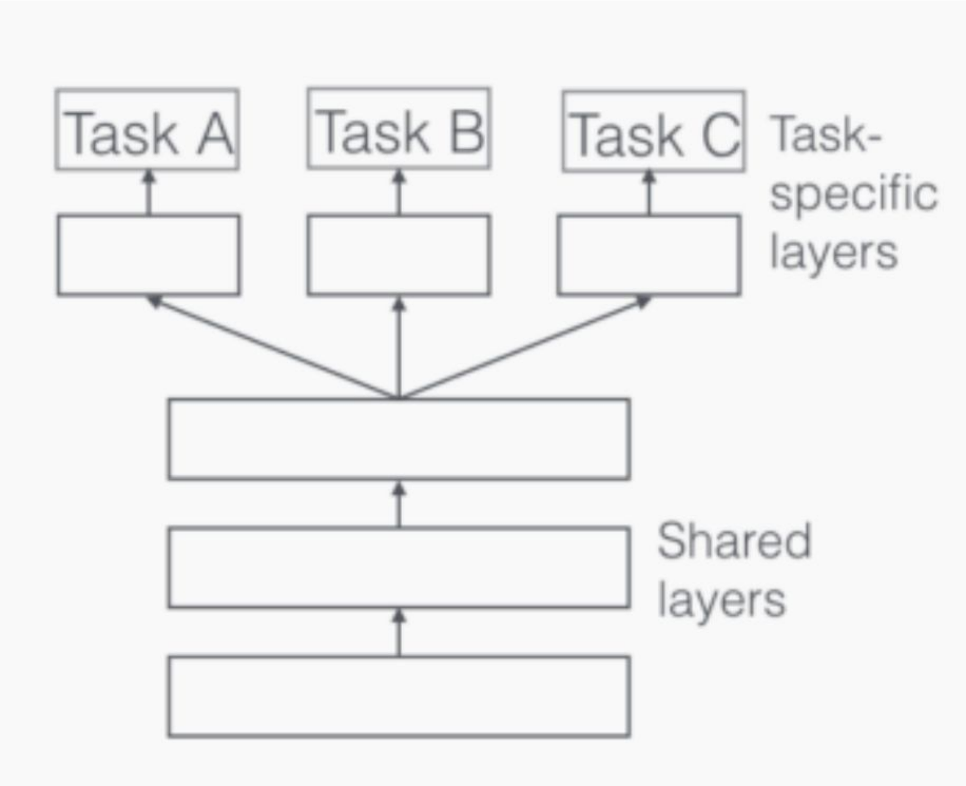




Transfer Learning with Pre-trained Deep Learning Models as Feature Extractors

Concept

Multi-Task Learning:



Concept

Relation:

Transfer Learning Settings	Related Areas	Source Domain Labels	Target Domain Labels	Tasks
Inductive Transfer Learning	Multi-task Learning	Available	Available	Regression, Classification
	Self-taught Learning	Unavailable	Available	Regression, Classification
Transductive Transfer Learning	Domain Adaptation, Sample Selection Bias, Co-variate Shift	Available	Unavailable	Regression, Classification
Unsupervised Transfer Learning		Unavailable	Unavailable	Clustering, Dimensionality Reduction

Outline

Motivation

Goals

Concept

Model Structure

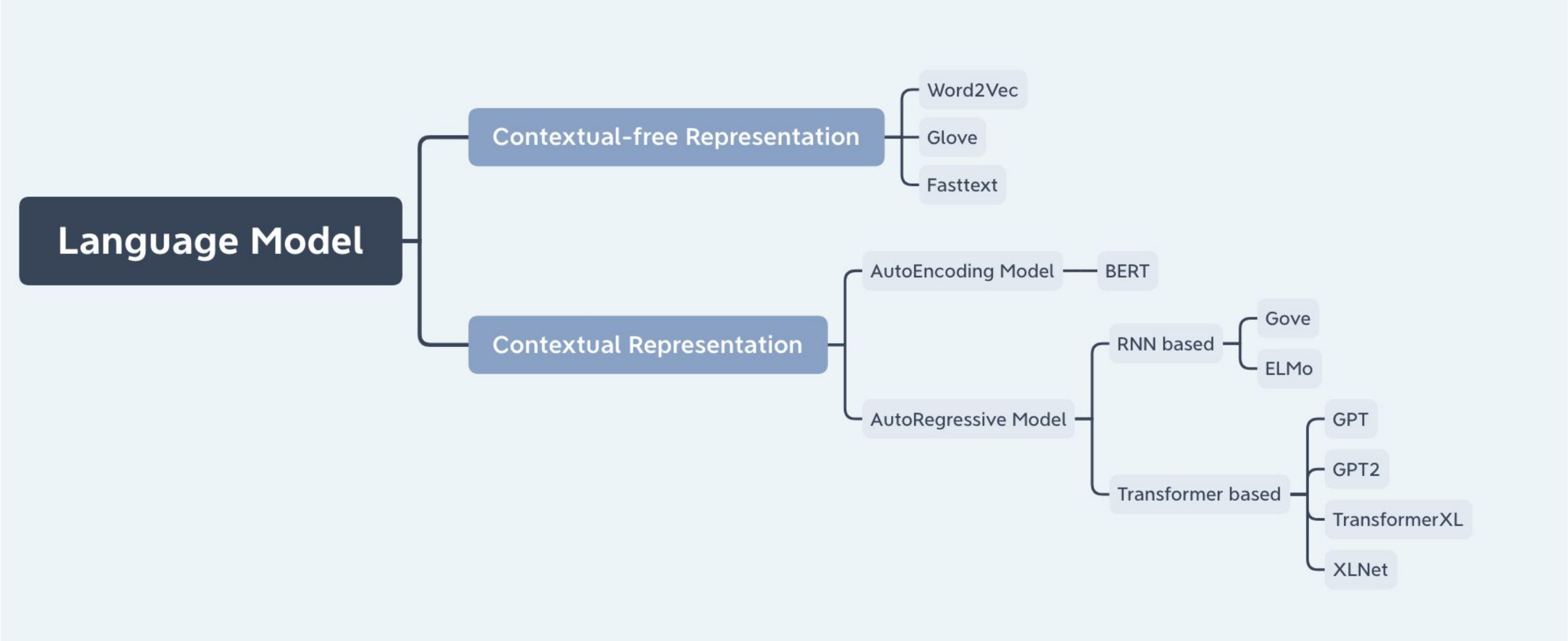
Tasks

- Text
- Protein
- Source code

Conclusion

Outlook

Model structure



Model principle

Contexual free models: Co-occurance Matrix

	C1	C2	C3	C4	C5
W1	1	0	0	2	0
W2	0	4	1	0	0
W3	2	0	0	1	0

AutoEncoding models: Masked tokens

RNN based models: Recurrent Neural Network

Transformer based model: $P(x) = \prod_{t=1}^{T} p(x_t|x_{<t})$ or $P(x) = \prod_{t=T}^{1} p(x_t|x_{>t})$.

Comparison Table

	Pros	Cons
Contextual free model	Low computing resource	only allow single context independent representation
AutoRegressive model--RNN	generate context dependent embeddings	Can't encode too long sentences
AutoEncoding model	supports bidirectional context reconstruction	inappropriate assumption: all masked tokens are constructed separately
AutoRegressive model--Transformer	consider sequential relationship between tokens	only encode a uni-directional context

Outline

Motivation

Goals

Concept

Model Structure

Tasks

- Text
- Protein
- Source code

Conclusion

Outlook

Tasks--TEXT

Task	Transfer Learning	MultiTask Learning	Model
Domain Adaption on Reading Comprehension	YES		BERT
Question Answering Sentence Selection	YES		BERT
Thermal dynamic modeling	YES		LSTM
Negation Detection	YES		BERT
Negation Detection		YES	BiLSTM
Natural language understanding		YES	BERT
Language Translation		YES	LSTM

Tasks--PROTEIN

Task	Transfer Learning	MultiTask Learning	Model
Protein structure prediction	YES		ELMo
Protein structure prediction	YES		Bi-direction Transformer
Protein structure prediction Protein Evolutionary Understanding	YES		LSTM
Protein ontology prediction Remote homology and fold prediction	YES		LSTM
Protein function prediction		YES	Multi-label Deep NN

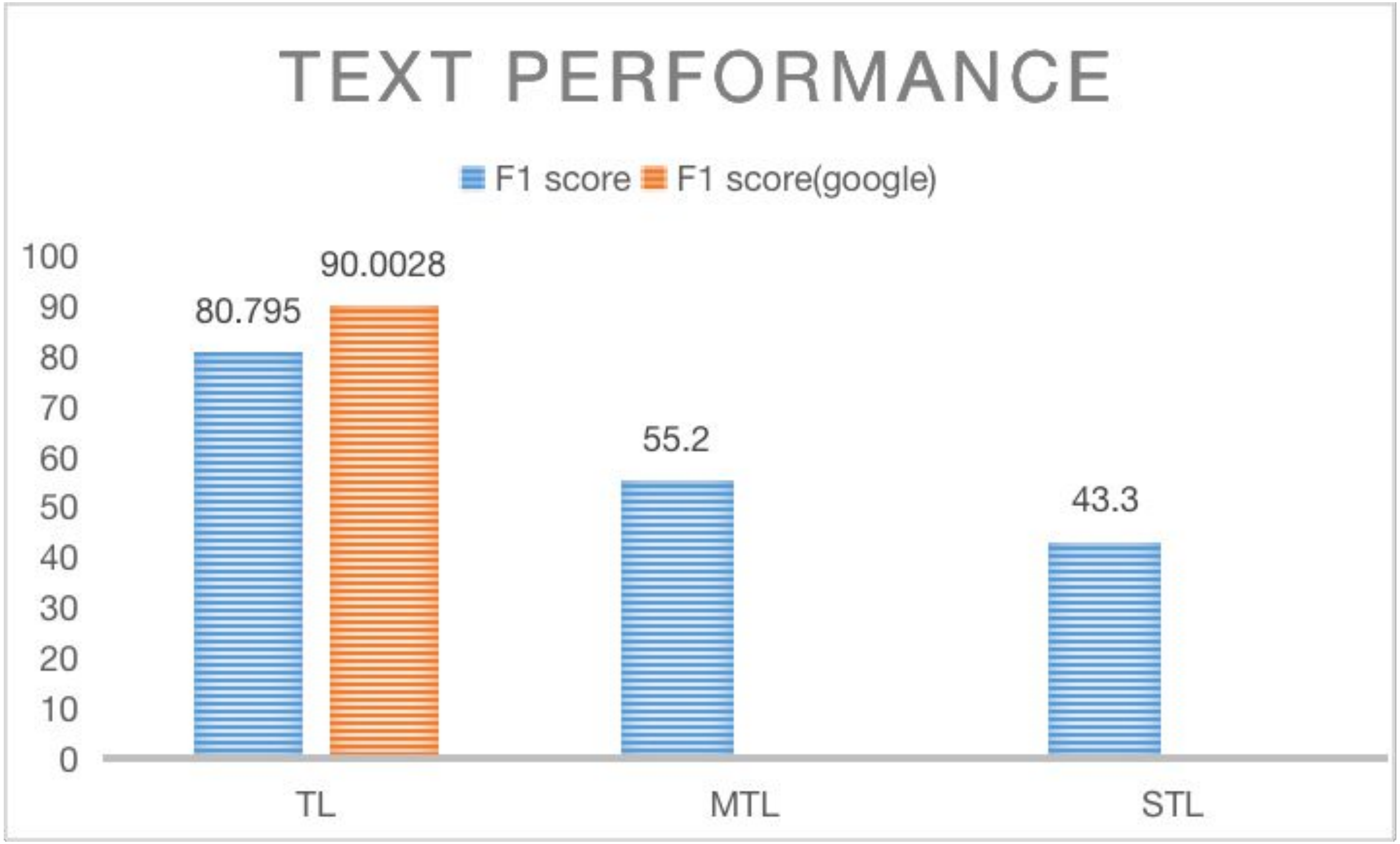
Tasks--Source Code

Task	Transfer Learning	MultiTask Learning	Model
Syntax detection	YES		LSTM
Code semantic embedding	YES		LSTM
Code clone detection	YES		RNN
API knowledge summarization Code summarization	YES		RNN
Semantic labeling	YES		CNN

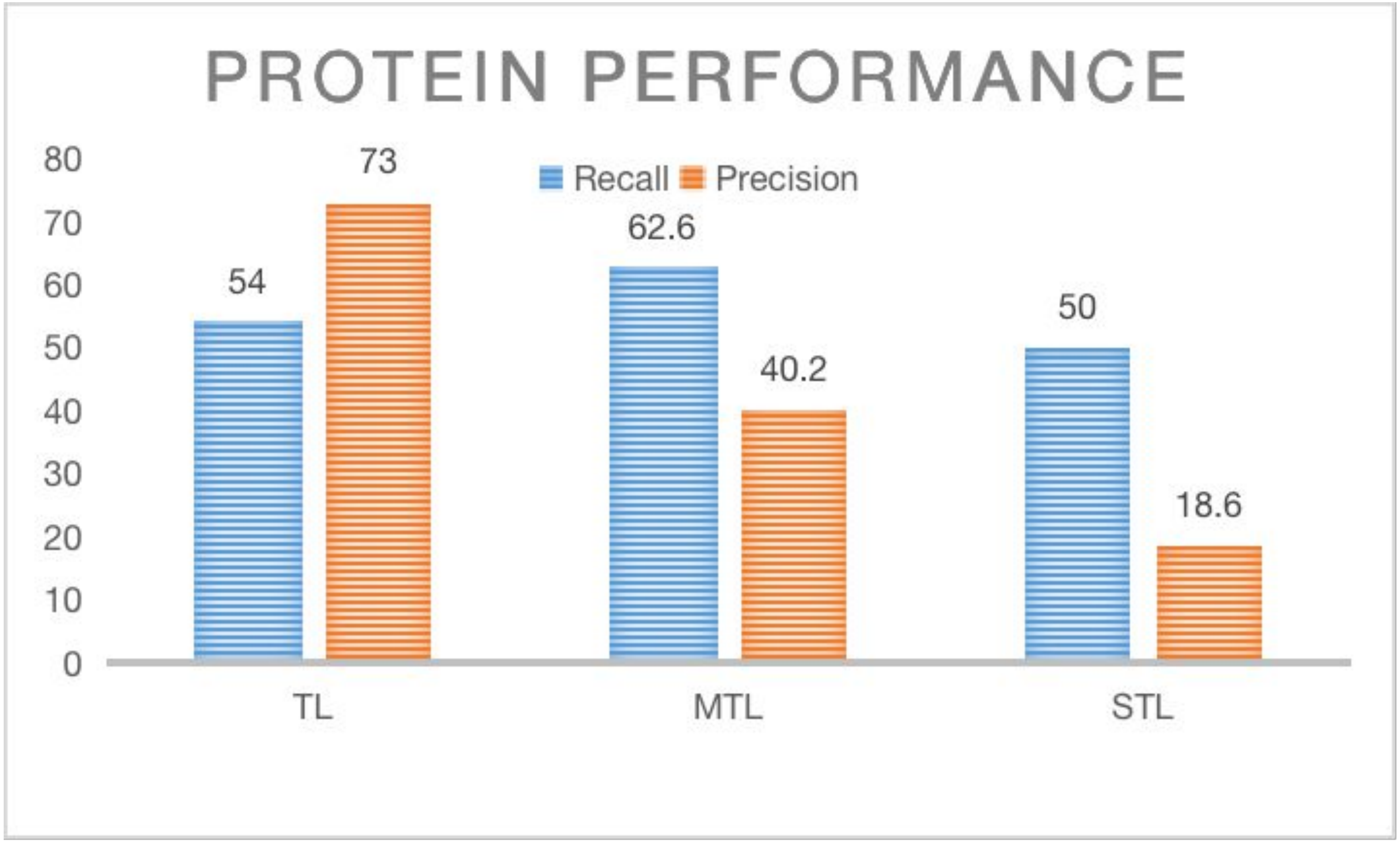
Pros and Cons

	Transfer Learning	MultiTask Learning
Pros	Make use of previous knowledge, no need to train from scratch. Low computing power	Introduce noisy data, increase generalization
Cons	Need massive pre-training data for generalization	Need to find suitable related tasks. Computing power needed

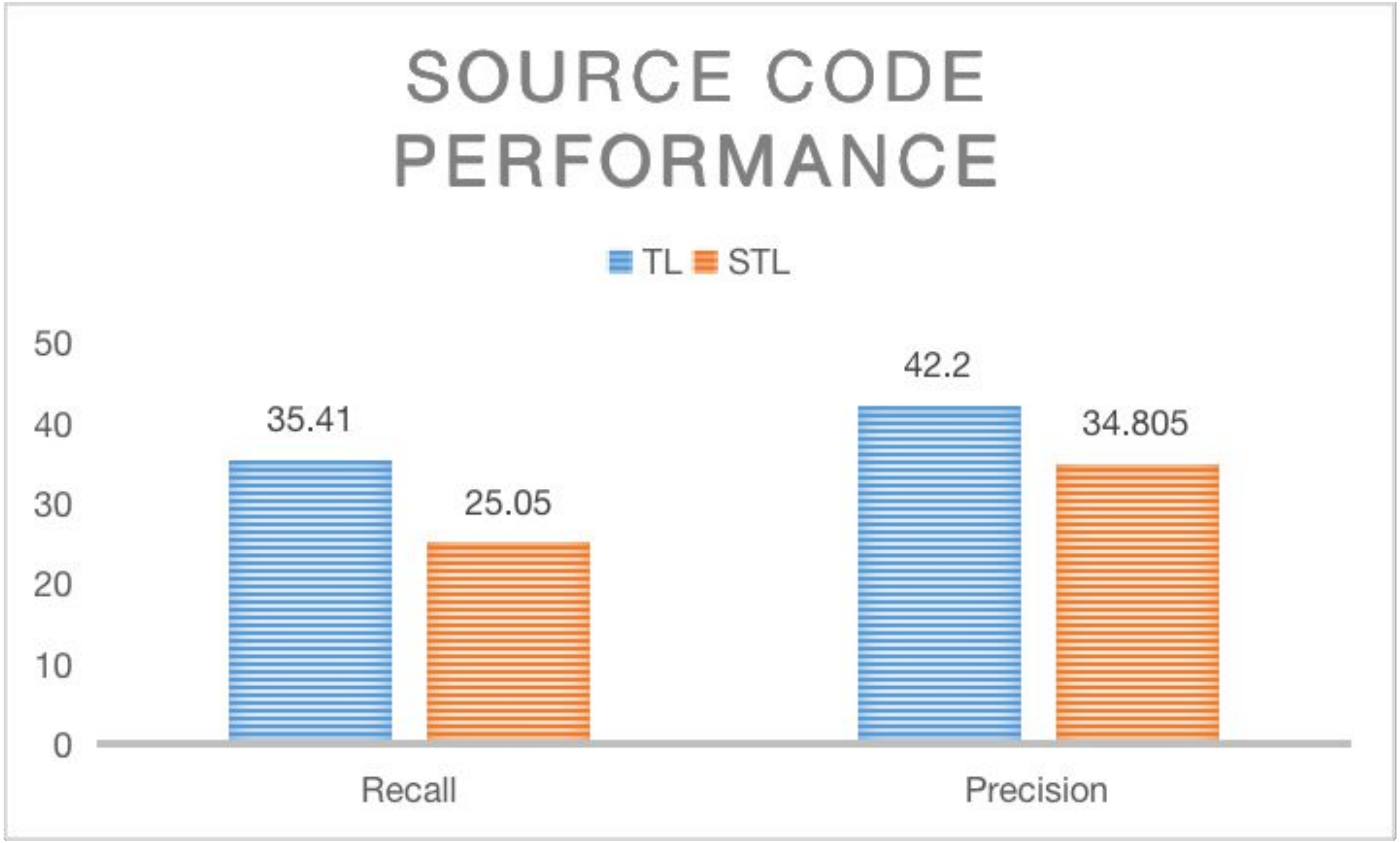
Statistics : Text



Statistics : Protein



Statistics : Source code



Outline

Motivation
Goals
Concept
Model Structure
Tasks
- Text
- Protein
- Source code

Conclusion
Outlook

Conclusion

Transfer Learning

Enough pre-training data

Multi-task Learning

Related tasks

Image source:https://cn.dreamstime.com/%E5%85%8D%E7%89%88%E7%A8%8E%E5%BA%93%E5%AD%98%E5%9B%BE%E7%89%87-3d%E5%B0%8F%E6%84%9F%E5%8F%B9%E5%8F%B7%E7%9A%84%E4%BA%BA-image12547789

Outline

Motivation

Goals

Concept

Model Structure

Tasks

- Text
- Protein
- Source code

Conclusion

Outlook

Outlook

Transfer Learning

Multi-task learning

Thanks for your attention !

Appendix1--TEXT

Task	F1	MAP	MAPE	Mean accuracy	Glue	Perplexity
Domain Adaption on Reading Comprehension	87.06					
Question Answering Sentence Selection		93.3				
Thermal dynamic modeling			1.396%			
Negation Detection	94.53					
Negation Detection				86.04		
Natural language understanding					93.1	
Language Translation						8.2

Appendix2--PROTEIN

Task	Q3	PLOT	Precision	Recall	AUC	F1
Protein structure prediction	70.3					
Protein structure prediction		PLOT				
Protein structure prediction Protein Evolutionary Understanding			73			
Protein ontology prediction Remote homology and fold prediction				54	89	
Protein function prediction			38.9	62.6		48

Appendix3--SOURCE CODE

Task	MRR	Accuracy	AUC	Precision	Recall	F-score
Syntax detection	0.52					
Code semantic embedding		88.09				
Code clone detection			82.4			
API knowledge summarization Code summarization				42.2	35.41	37.91
Semantic Labeling			0.769			