

# 文本情感分类

导师：罗老师

---



# Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks

从树结构的长短期记忆网络改进语义表示

作者: Kai Sheng Tai

单位: Stanford University

发表会议及时间: ACL 2015年



# 前期知识储备

Pre-knowledge reserve



## Pytorch

基本熟悉Pytorch环境搭建以及基本代码结构

## RNN

了解循环神经网络（RNN）的结构，掌握RNN的基本工作原理

## LSTM

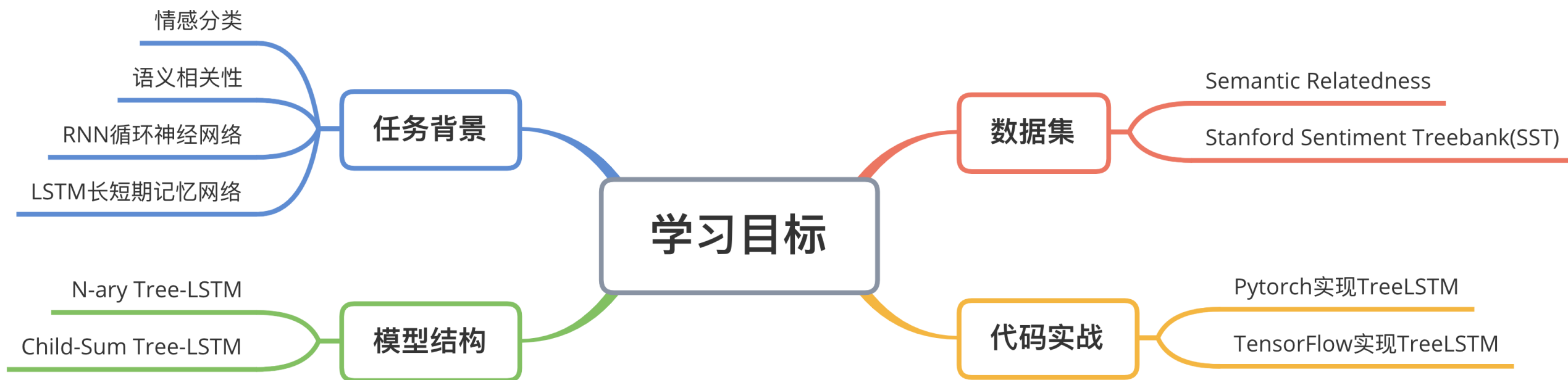
了解长短期记忆网络（LSTM）的结构，掌握LSTM的基本工作原理

## 词向量使用

了解词向量的训练方法，包括Word2Vec, Glove

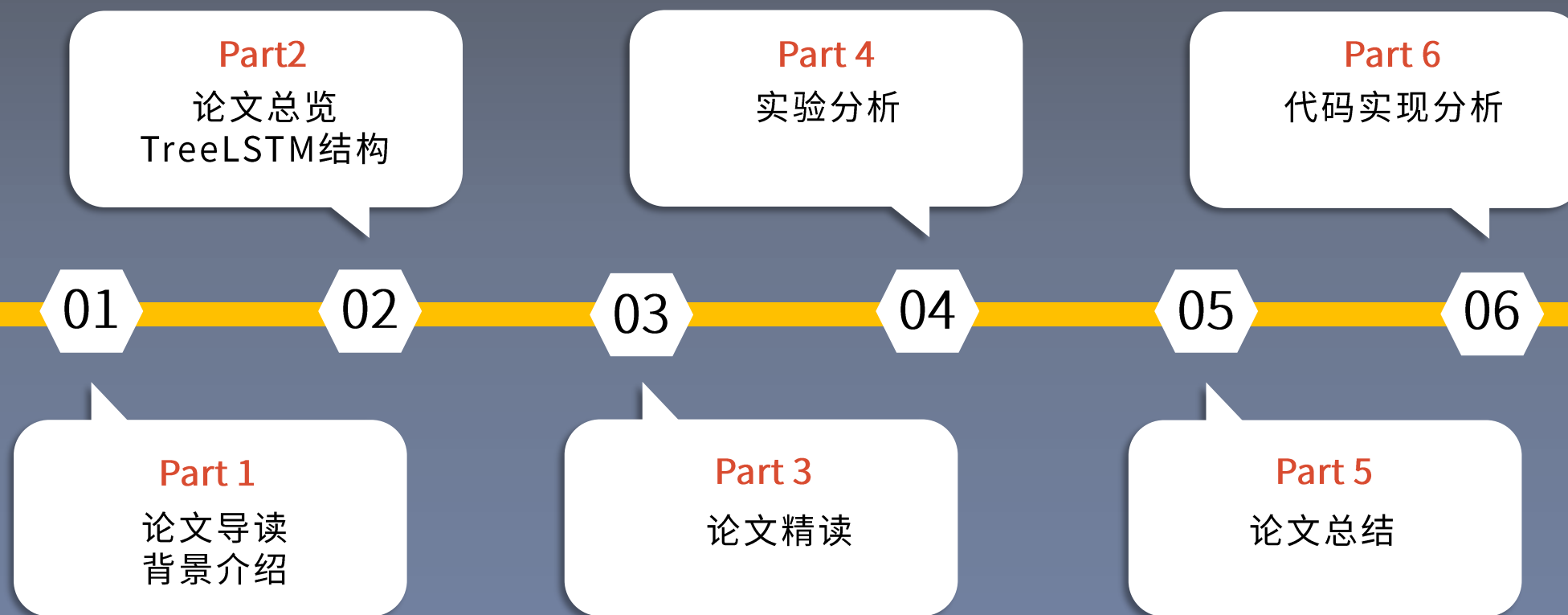
# 学习目标

Learning objectives



# 课程安排

The schedule of course





# 第一课：论文导读

The first lesson: the paper guide

---



# 目 录

1/ 论文研究背景、成果及意义

2/ 论文泛读

3/ 本课回顾及下节预告



# 论文研究背景、成果及意义

---



# 研究背景

Research background

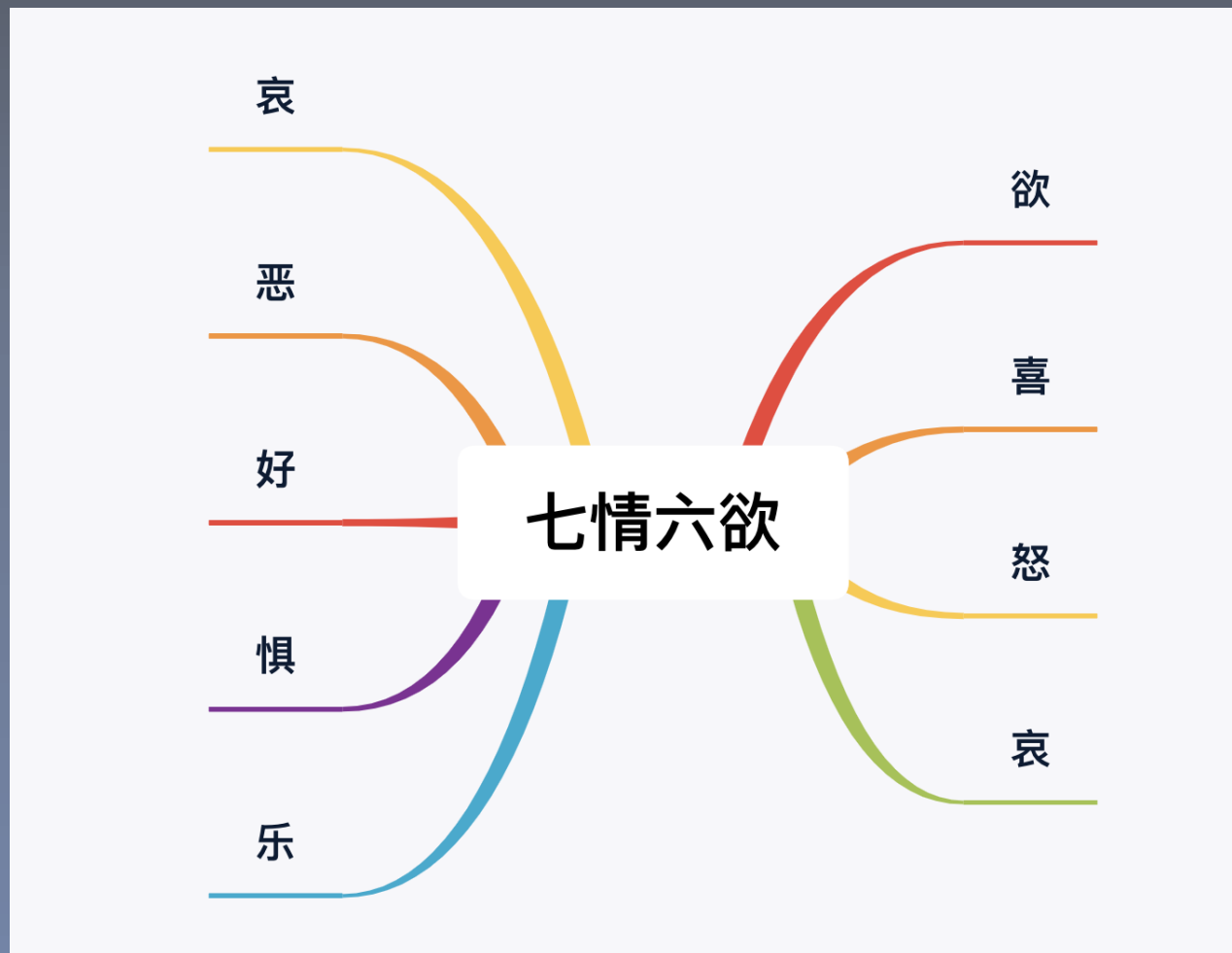


深度之眼  
deepshare.net



重点 重点来了!

## 情感分类



# 研究背景

Research background

特喜欢这种好看的狗狗



积极

90%



10%



消极

这真是惊艳世界的中国“黑科技”



积极

65%



35%



消极

这车悬挂特别硬，完全难以忍受



积极

6%



94%



消极

环境特别差，脏兮兮的，再也不去了



积极

1%



99%



消极



# 研究背景

Research background

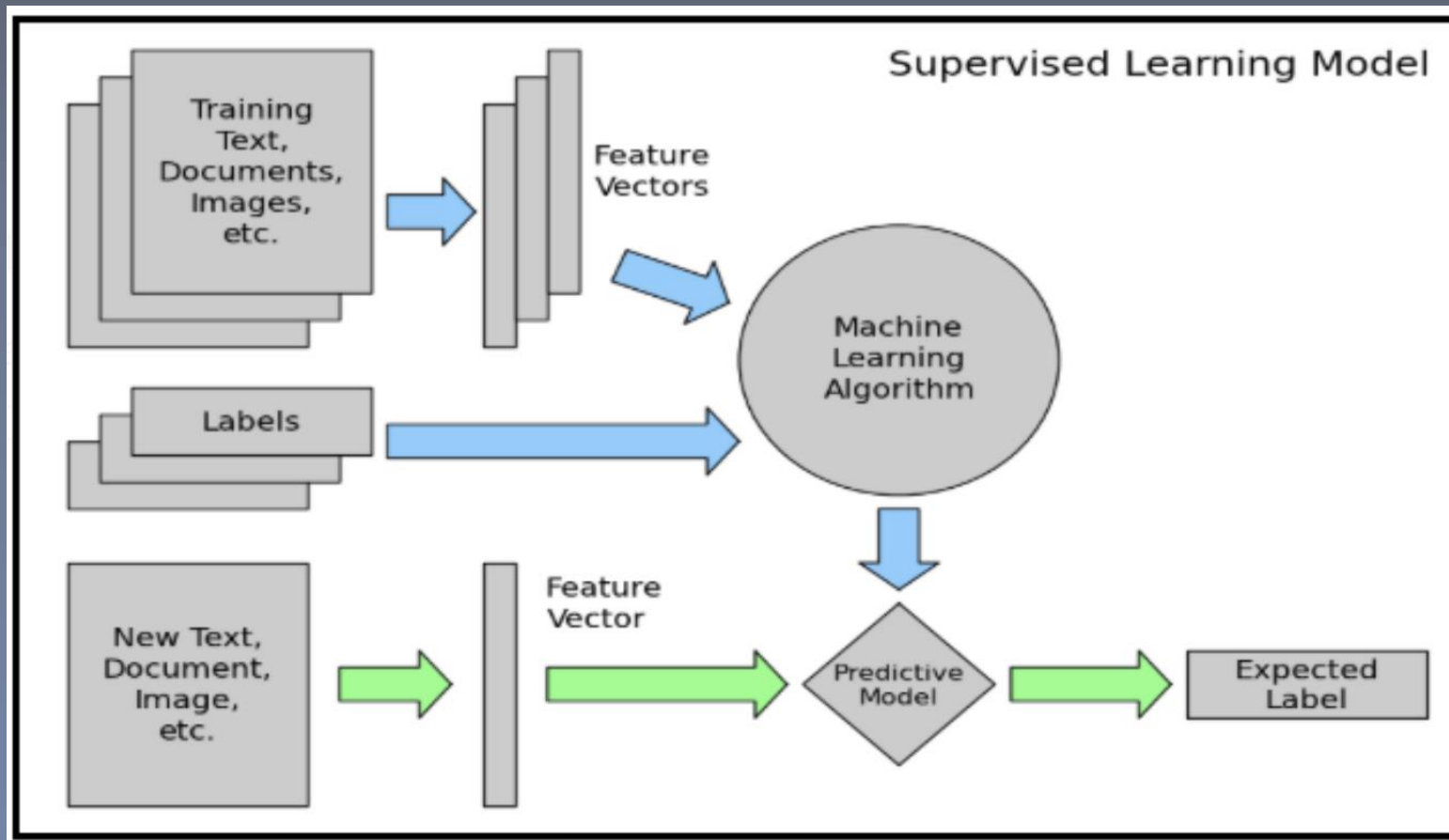
## 监督学习模型



深度之眼  
deepshare.net



重点 重点来了!





# 研究背景

Research background

## 循环神经网络

解决任意长度  
句子分类问题

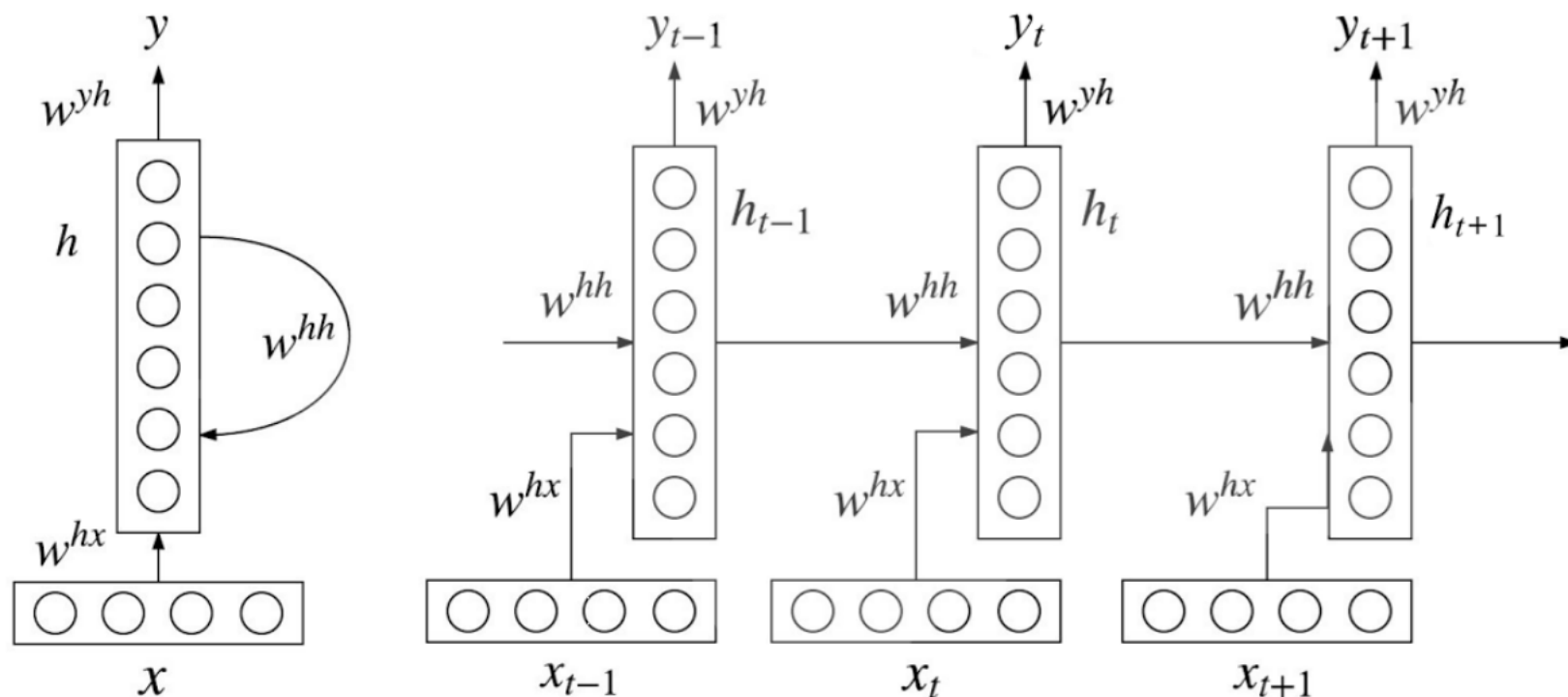


Figure 5: Recurrent neural network



# 研究背景

Research background



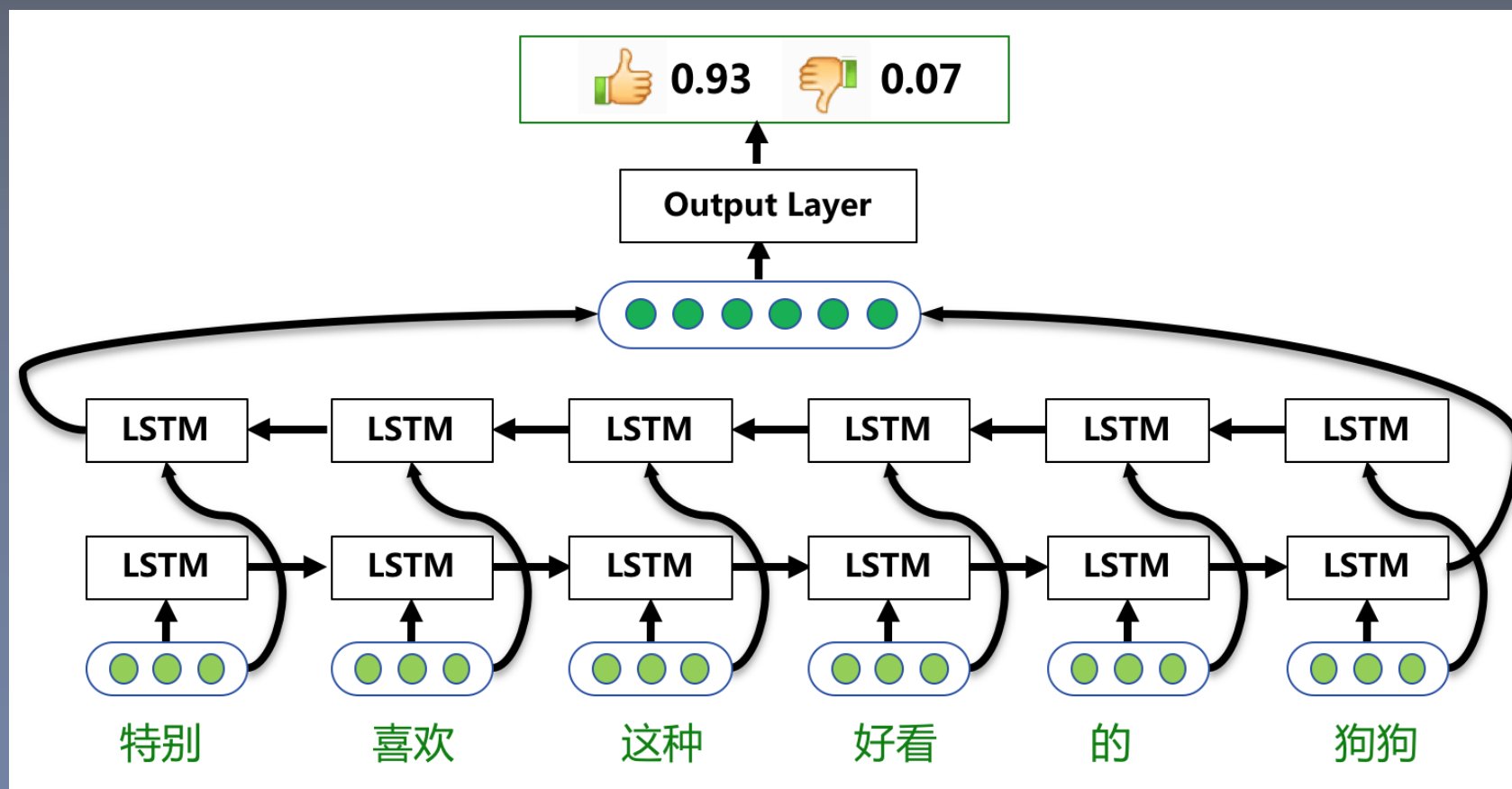
深度之眼  
deepshare.net



重点 重点来了!

## 长短期记忆网络

一定程度上降低  
梯度消失和  
爆炸的程度



# 研究背景

Research background

---



深度之眼  
deepshare.net



重点 重点来了!

## 如何有效的学习语法语境是 否和句子结构的关系?



# 研究背景

Research background

## 语法分析



重点 重点来了!

语言其实具备自顶而下的层级关系，固定数量的语法结构能够生成无数句子。比如，仅仅利用下列两个语法规律，我们就能够生成所有名词短语。

- 名词短语可以由名词和名词短语组成。
- 名词短语还可以由名词和名词组成。

上海浦东机场航站楼

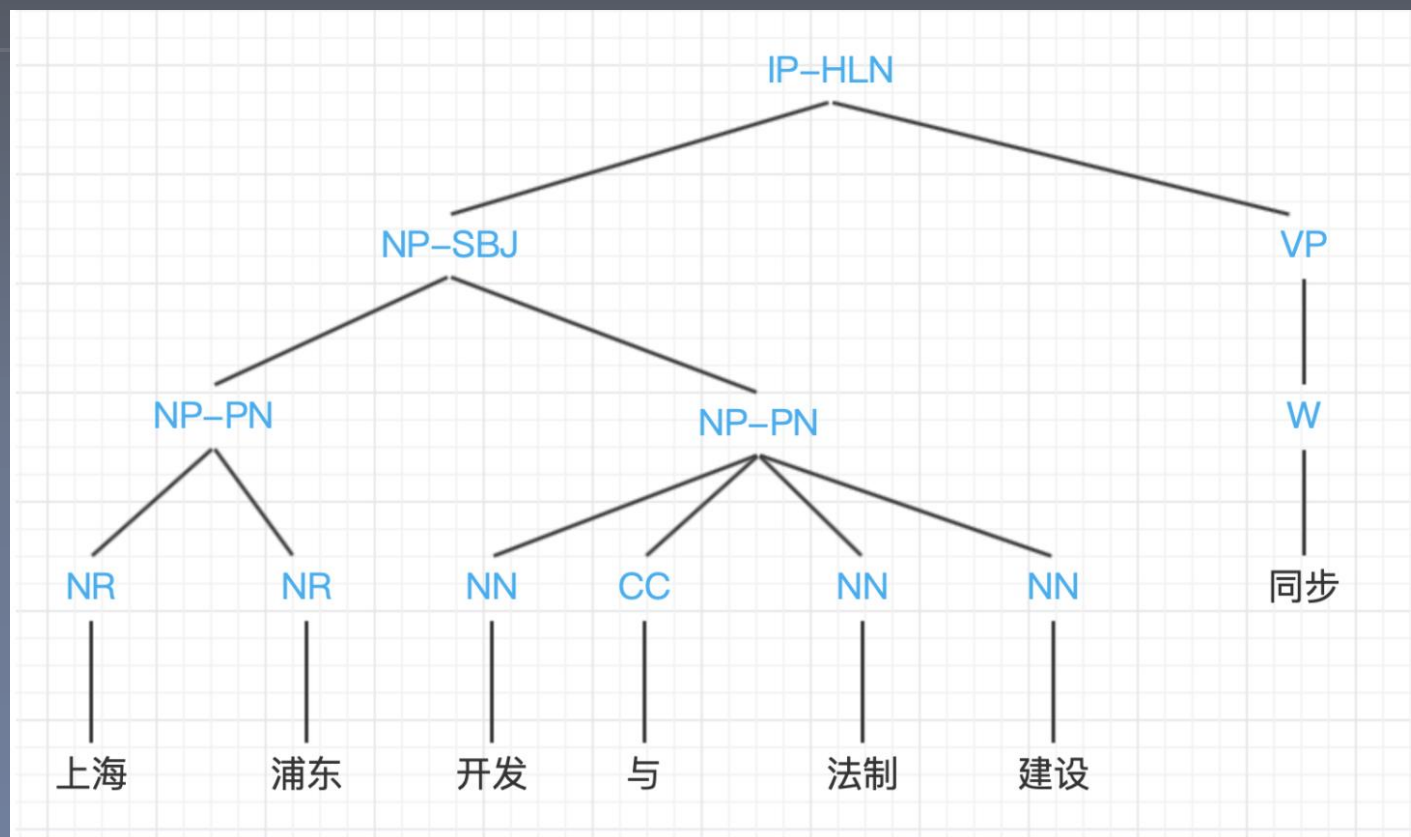
上海 + 浦东 + 机场 + 航站楼



# 研究背景

Research background

## 短语结构树



重点 重点来了!

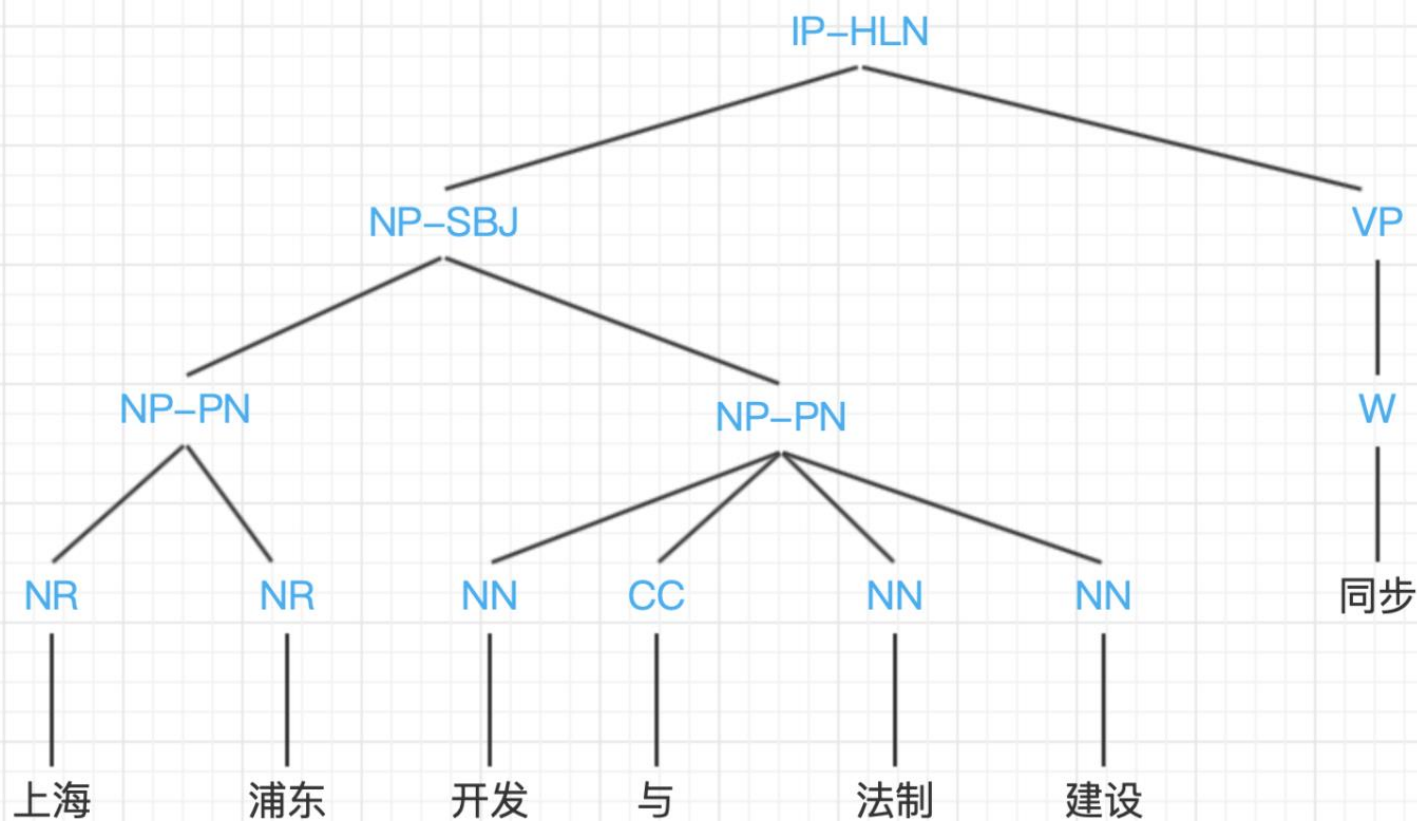
短语结构语法描述了如何自顶而下的生成一个句子，反过来，句子也可以用短语结构语法来递归的分解。层级结构其实是一种树形结构，例如这句话“上海 浦东 开发与 法制 建设 同步”，分解成如下图的短语结构树：





# 研究背景

Research background

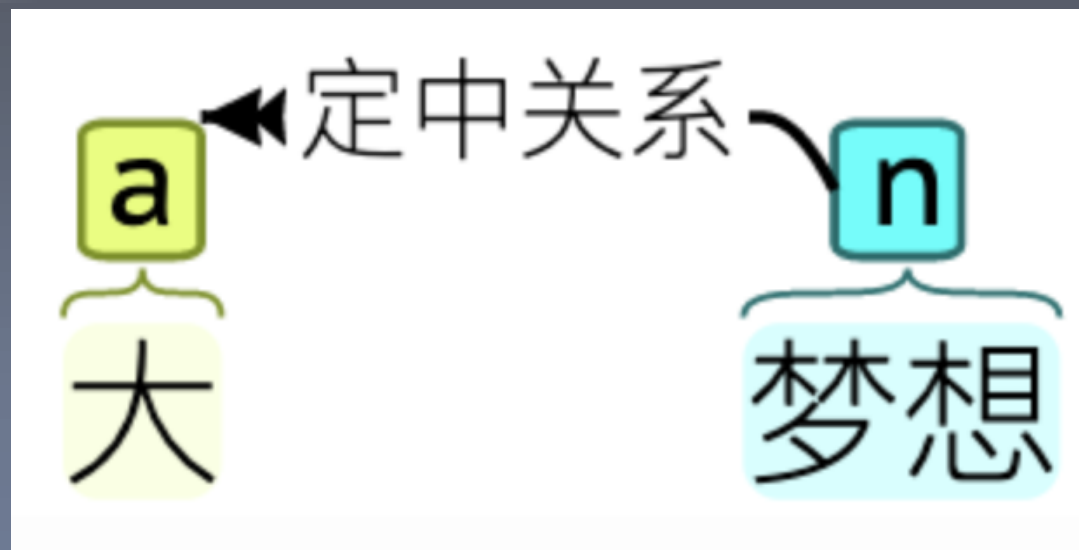


标记	释义
IP-HLN	单句-标题
NP-SBJ	名词短语-主语
NP-PN	名词短语-代词
NP	名词短语
VP	动词短语

# 研究背景

Research background

## 依存句法树



重点 重点来了!

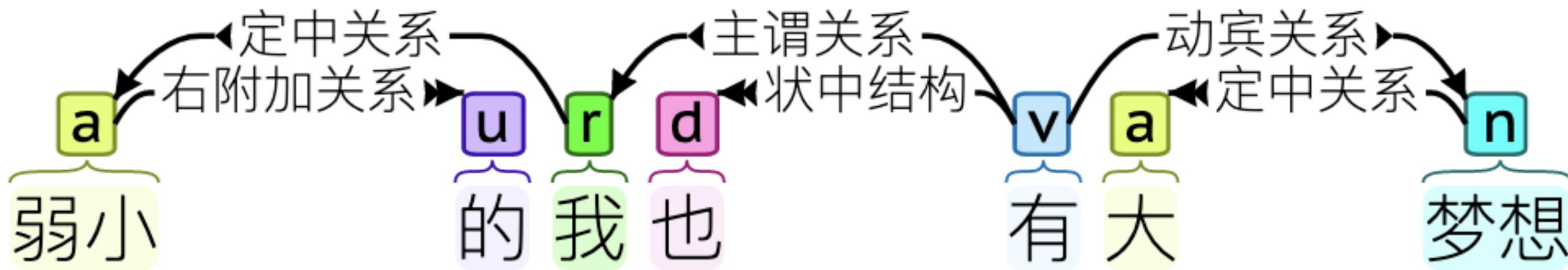
依存语法理论认为词与词之间存在主从关系，这是一种二元不等价的关系。在句子中，如果一个词修饰另一个词，则称修饰词为从属词( dependent )，被修饰的词语称为支配词(head),两者之间的语法关系称为依存关系( dependency relation)。比如句子“大梦想”中形容词“大”与名词“梦想”之间的依存关系





# 研究背景

Research background



现代依存语法中，语言学家 Robinson 对依存句法树提了 4 个约束性的公理。

- 有且只有一个词语(ROOT，虚拟根节点，简称虚根)不依存于其他词语。
- 除此之外所有单词必须依存于其他单词。
- 每个单词不能依存于多个单词。
- 如果单词 A 依存于 B，那么位置处于 A 和 B 之间的单词 C 只能依存于 A、B 或 AB 之间的单词。

这 4 条公理分别约束了依存句法树(图的特例)的根节点唯一性、连通、无环和投射性( projective )。这些约束对语料库的标注以及依存句法分析器的设计奠定了基础。

# 研究背景

Research background

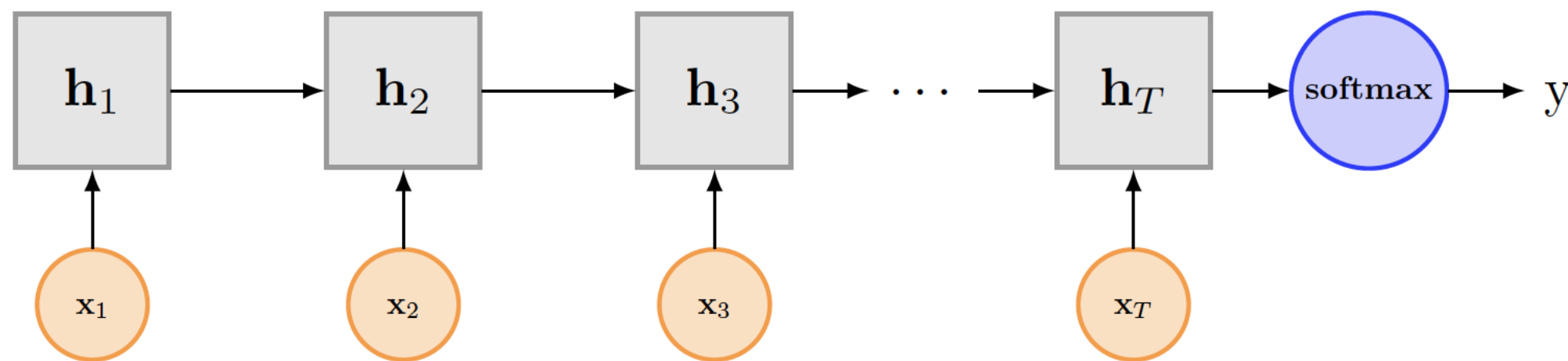


Figure 1: Recurrent Neural Network for Classification

# 研究背景

Research background

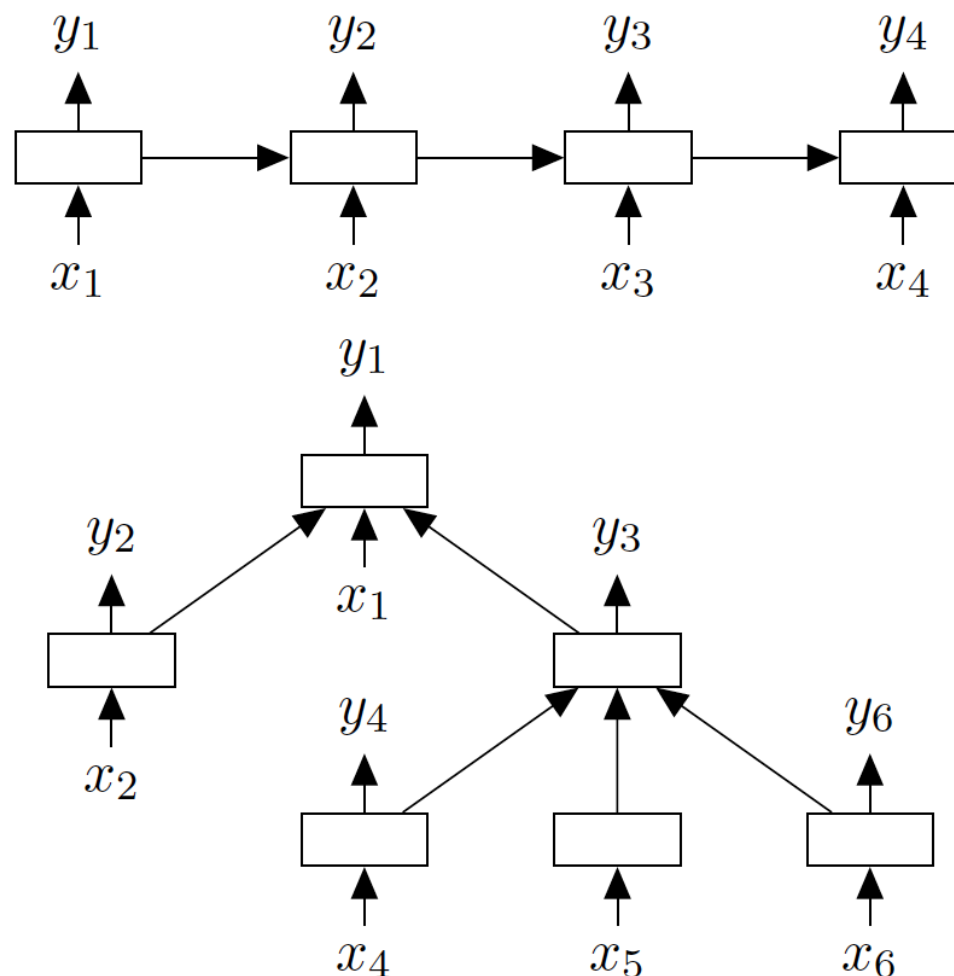


Figure 1: **Top:** A chain-structured LSTM network. **Bottom:** A tree-structured LSTM network with arbitrary branching factor.



# 论文泛读

Strcuture of Paper

---



# 论文结构

## Structure of Papers



# 摘要

abstract

---

## 摘要核心

1. 论文的主要工作就是丰富了LSTM的拓扑结构，由我们所熟悉的线性LSTM结构衍生出了树形LSTM结构，从而更方便地结合依存关系、短语构成等语法特性，使得语义表达更加准确。论文在预测两个句子的语义相关性(SICK) 和情感分类 (Stanford Sentiment Treebank)两个任务上做了实验，验证了新LSTM模型的效果。



# 论文小标题

Paper title

---

1. Introduction
2. Long Short-Term Memory Networks
3. Tree-Structured LSTMs
  - 3.1 Child-Sum Tree-LSTMs
  - 3.2 N-ary Tree-LSTMs
4. Models
  - 4.1 Tree-LSTM Classification
  - 4.2 Semantic Relatedness of Sentence Pairs

5. Experiment
  - 5.1 Sentiment Classification
  - 5.2 Semantic Relatedness
  - 5.3 Hyperparameters and Training Details
6. Results
  - 6.1 Sentiment Classification
  - 6.2 Semantic Relatedness
7. Discussion and Qualitative Analysis
  - 7.1 Modeling Semantic Relatedness
  - 7.2 Effect of Sentence Length
8. Related Work
9. Conclusion



# 本课回顾及下节预告

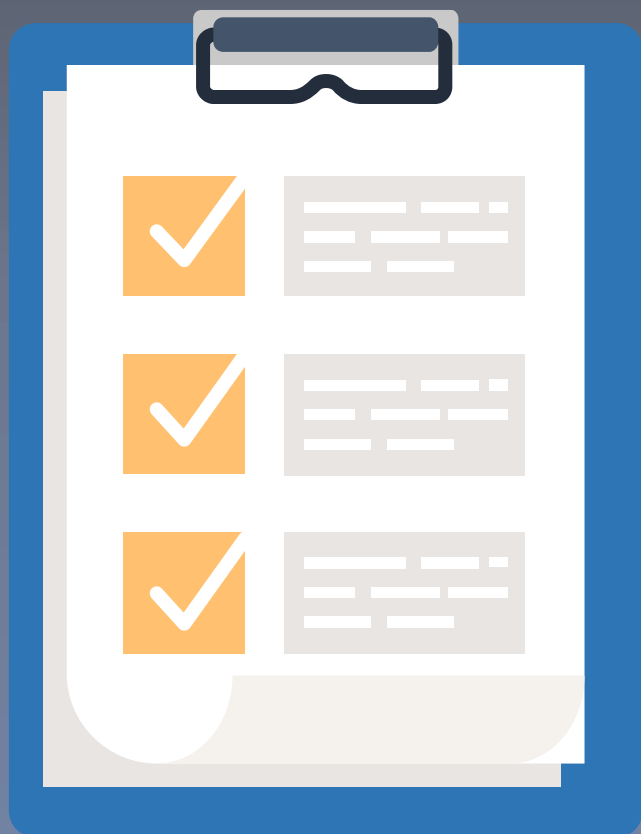
Review in the lesson and Preview of next lesson

---



# 本课回顾

Review in the lesson



## 01 背景介绍

简单回顾情感分类任务

## 02 循环神经网络的情感分析任务

介绍基于RNN/LSTM的循环神经网络结构

## 03 语法分析任务以及相关任务背景

详细介绍了语法分析的作用，从文本结构上带来的理解

## 04 论文泛读

泛读结构，了解论文大体结构



# 下节预告

Preview of next lesson



## 01 循环神经网络的详细讲解

介绍RNN/LSTM的结构以及优缺点

## 02 基于TreeLSTM的情感分类模型

精读论文，学习模型结构

## 03 实验设置及结果分析

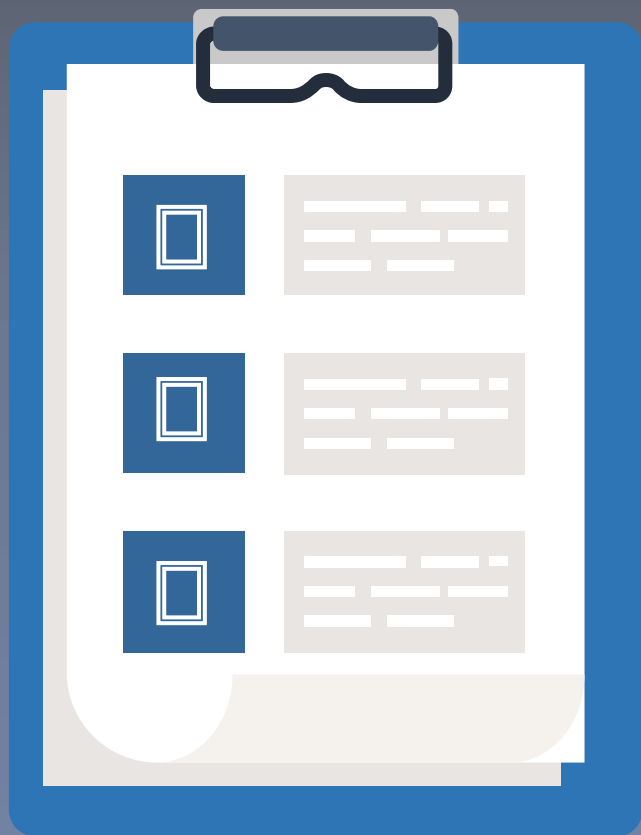
网络超参数设置，学习率，batchsize等  
实验结果分析对比

## 04 论文总结

总结论文中创新点、关键点及启发点

# 下节课前准备

Preview of next lesson



- 下载论文
- 泛读论文
- 筛选出自己不懂的部分，带着问题进入下一课时

# —— 结 语 ——

在其他任何地方的投资，都没有投资自己回报率高，你是赚不到超出你认知以外的钱的

愿你成为一名终身学习者。

如果有一天学累了，坚持不下去了，告诉自己，如果我现在觉得学习痛苦，是因为我还不知道无知的代价更大







深度之眼  
deepshare.net

联系我们：

电话：18001992849

邮箱：service@deepshare.net

QQ：2677693114



公众号



客服微信

