

Beijing Forest Studio  
北京理工大学信息系统及安全对抗实验中心



# 多视角深度学习

多视角深度学习

硕士研究生 董勃

2021年07月18日

- 预期收获
  - 1. 理解多视角的基本概念
  - 2. 了解多视角的应用条件
  - 3. 熟悉多视角的常见应用
  - 4. 了解多视角的优劣之处
  - 5. 利用多视角的思维解决专业问题

- 背景

- 大多数科学数据分析问题中，需要从**不同领域**或从**各种特征**中获得异构数据
- 不同类型数据可以自然地划分为组，每个变量组被称为特定的视角

- 基本概念

- 中文含义：从各种**不同的角度**进行观察同一事物
- 深度学习含义：引入一个函数去模型化一个特定的视角，并且利用相同输入的冗余视角去联合优化所有函数

- 多视角来源

- 多源：人物识别可以用脸、指纹等作为不同源的输入
- 多特征子集：图像表示可以用颜色、文字作为不同的特征表述



- 算法分类

- 协同训练：在未标记数据的两个不同视角下，轮流训练，使相互一致性最大化
- 多核学习：利用对应不同视角的多个内核，并线性或非线性地组合内核以提高学习性能
- 子空间学习：假设当前视角来源于某个潜在空间，来获取多个视角共享的潜在子空间

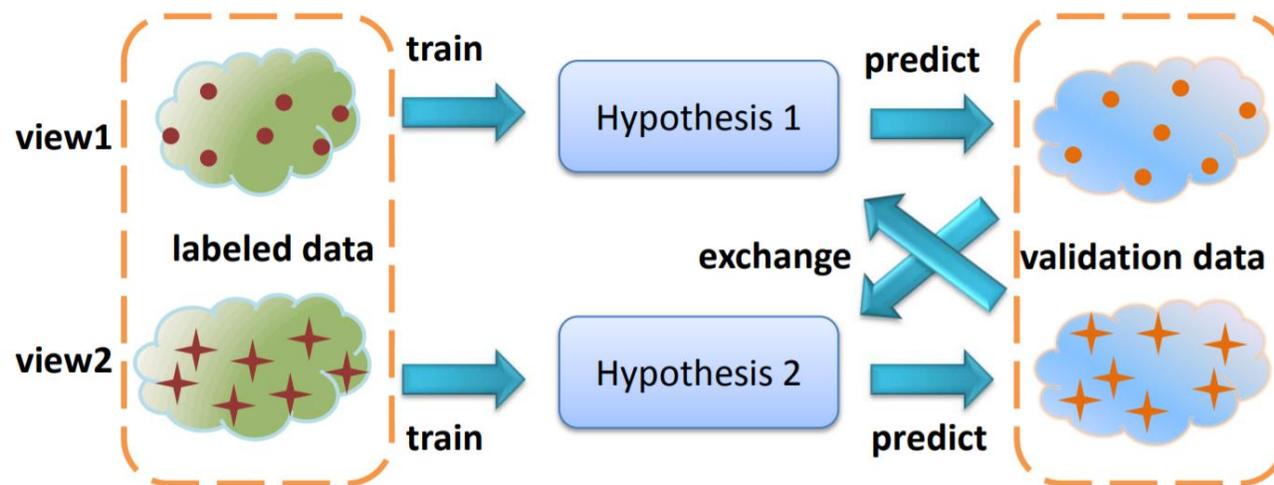
- 算法原则

- 共识准则  
    尽量保证先验知识一致
- 互补准则  
    每个视角都拥有其他视角没有的信息



- 协同训练

- 在未标记数据的两个不同视角下，轮流训练，使相互一致性最大化

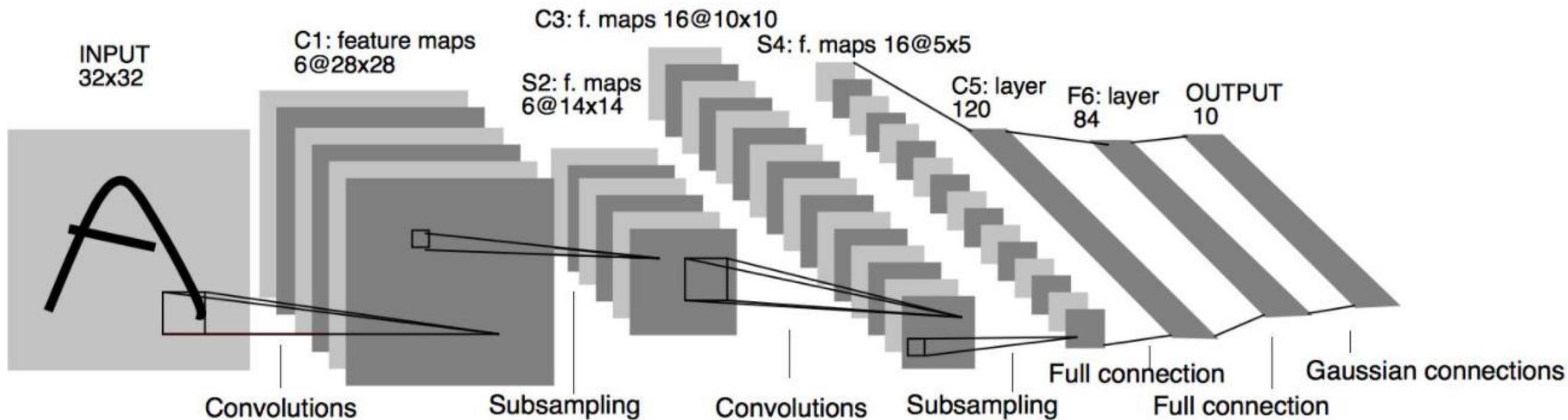


1. 已标注数据集L1，训练得到分类器F1；
2. 已标注数据集L2，训练得到分类器F2；
3. F1对未标记数据U进行分类，取前K个置信度最高的数据及标签加入L2；
4. F2对未标记数据U进行分类，取前K个置信度最高的数据及标签加入L1；
5. 重复步骤1-4，直到U为空集。



基本思路

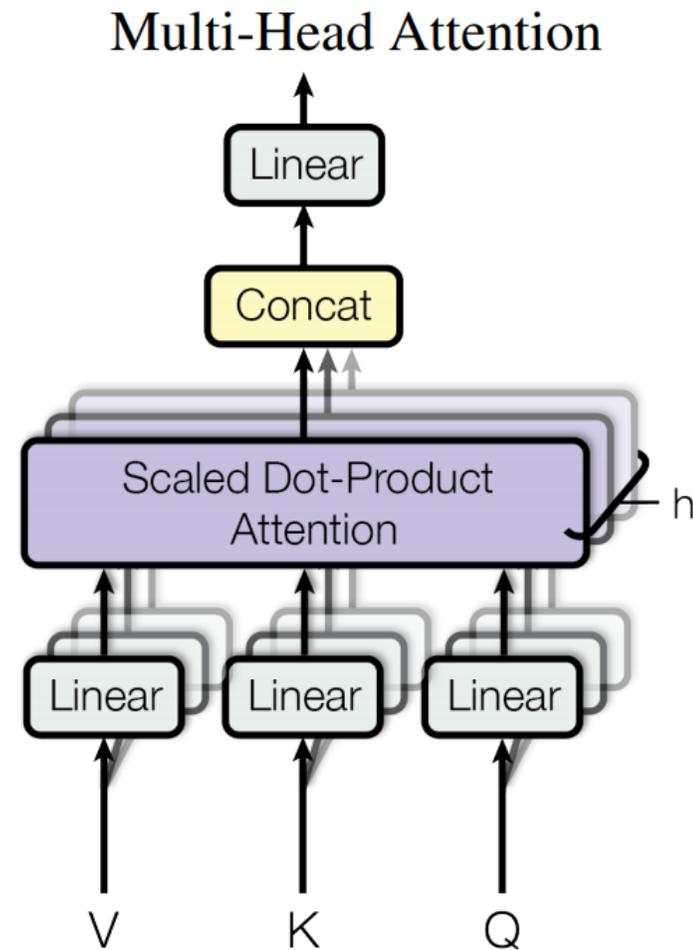
- 卷积核作用：局部感知野，信息混合
- 多卷积核
  - 卷积核的权重随机初始化
  - 不同卷积核学到不同的特征



- 多头自注意力机制
  - 做法：多个独立self-attention计算，然后进行集成
  - 作用：有助于网络从不同的特征子空间捕获更丰富的特征

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O$$

where  $\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$





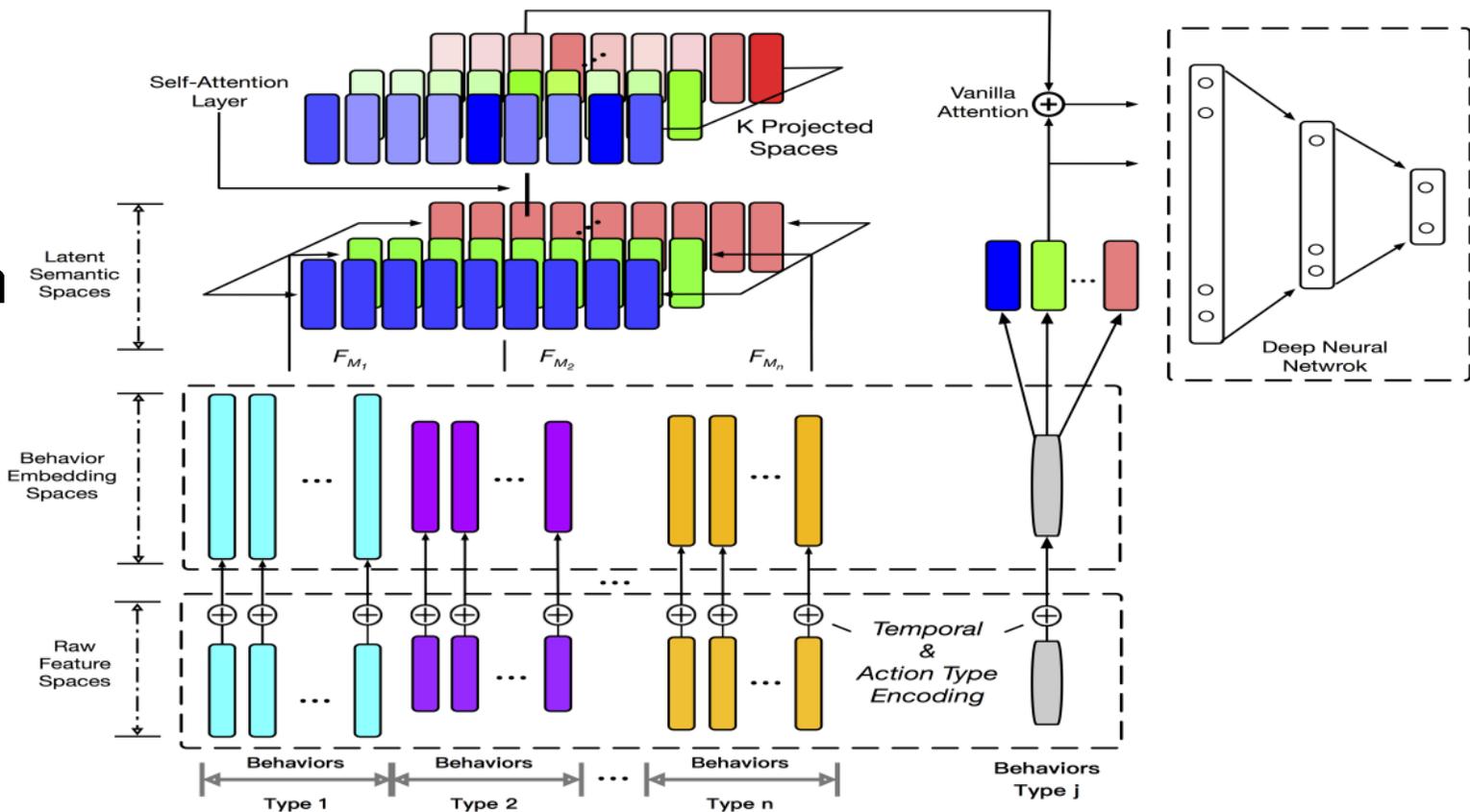
算法原理

T	针对用户异构行为，提出通用框架
I	多种类型用户行为
P	For { 1. 用户行为嵌入 2. 行为映射到多语义空间 3. attention建模行为关联 }
O	通用框架，用于各类下游任务

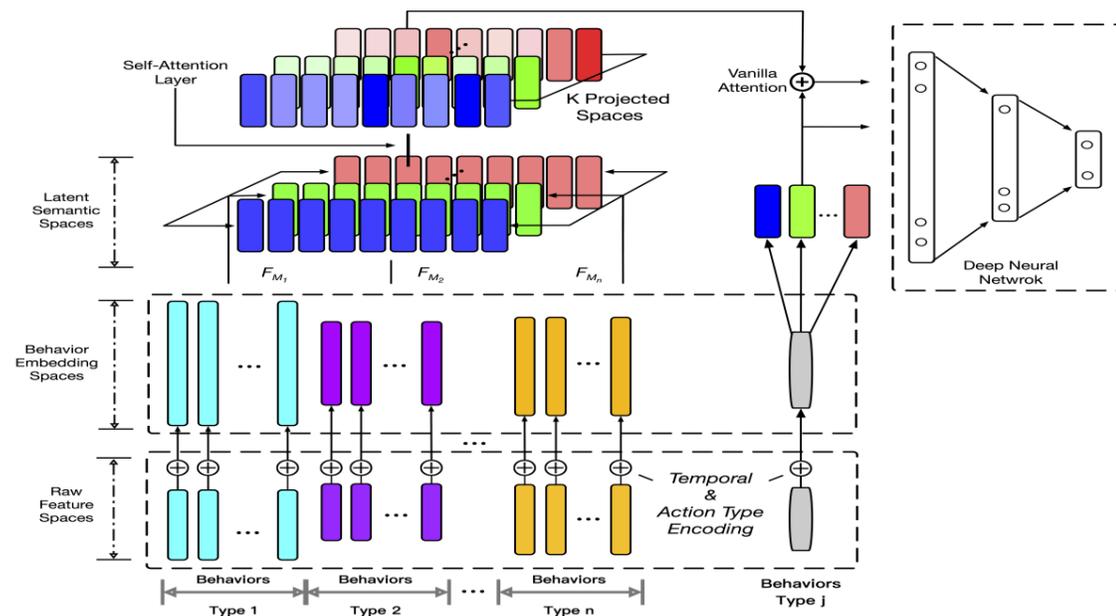
P	对用户异构行为数据进行更精细的处理
C	构建具备多种类型的用户行为数据
D	融合不同类型用户行为序列，并提出通用用户表征框架
L	AAAI2018

- 逻辑框架

- 原始特征层
- 语义映射层
- Self-attention层  
multi-head attention
- 目标网络  
下游任务



- 语义映射层
  - 做法：将异构行为线性映射到多个语义空间
  - 原因：让不同的行为在不同的语义空间下进行比较和相互作用
  - 解释：如果不分多语义空间，会发生所谓**语义中和现象**，导致不同种类的行为在不相关的领域上增大互相影响
  - 实现：行为编码向一个统一的空间映射

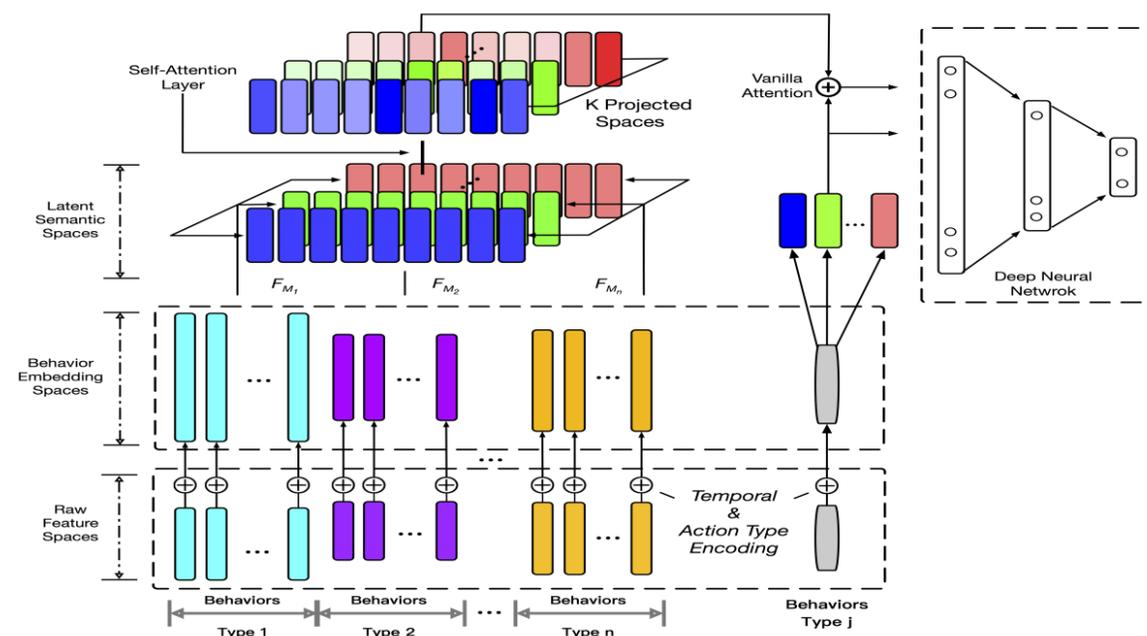


- Self Attention层
  - 作用：将用户的每一个行为从一个客观的表征，转换为用户记忆中的表征

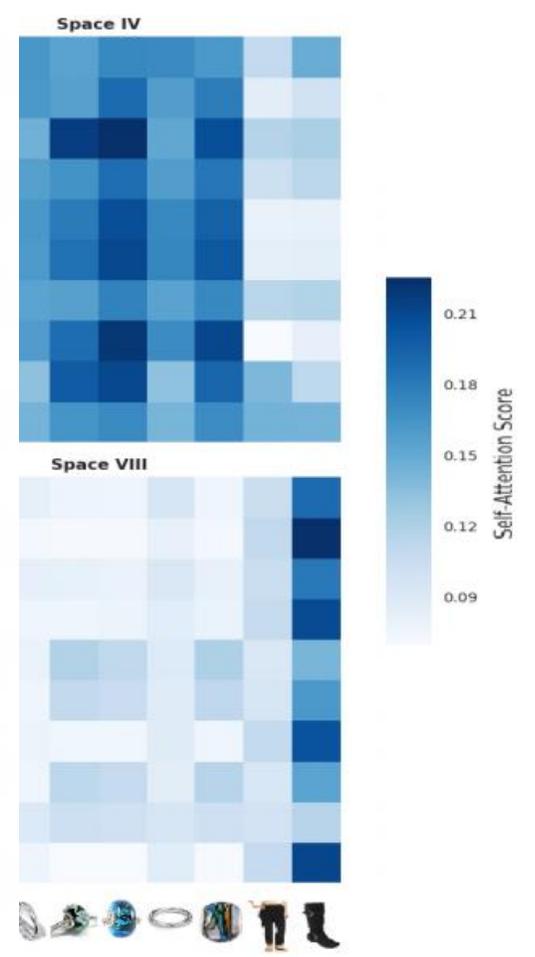
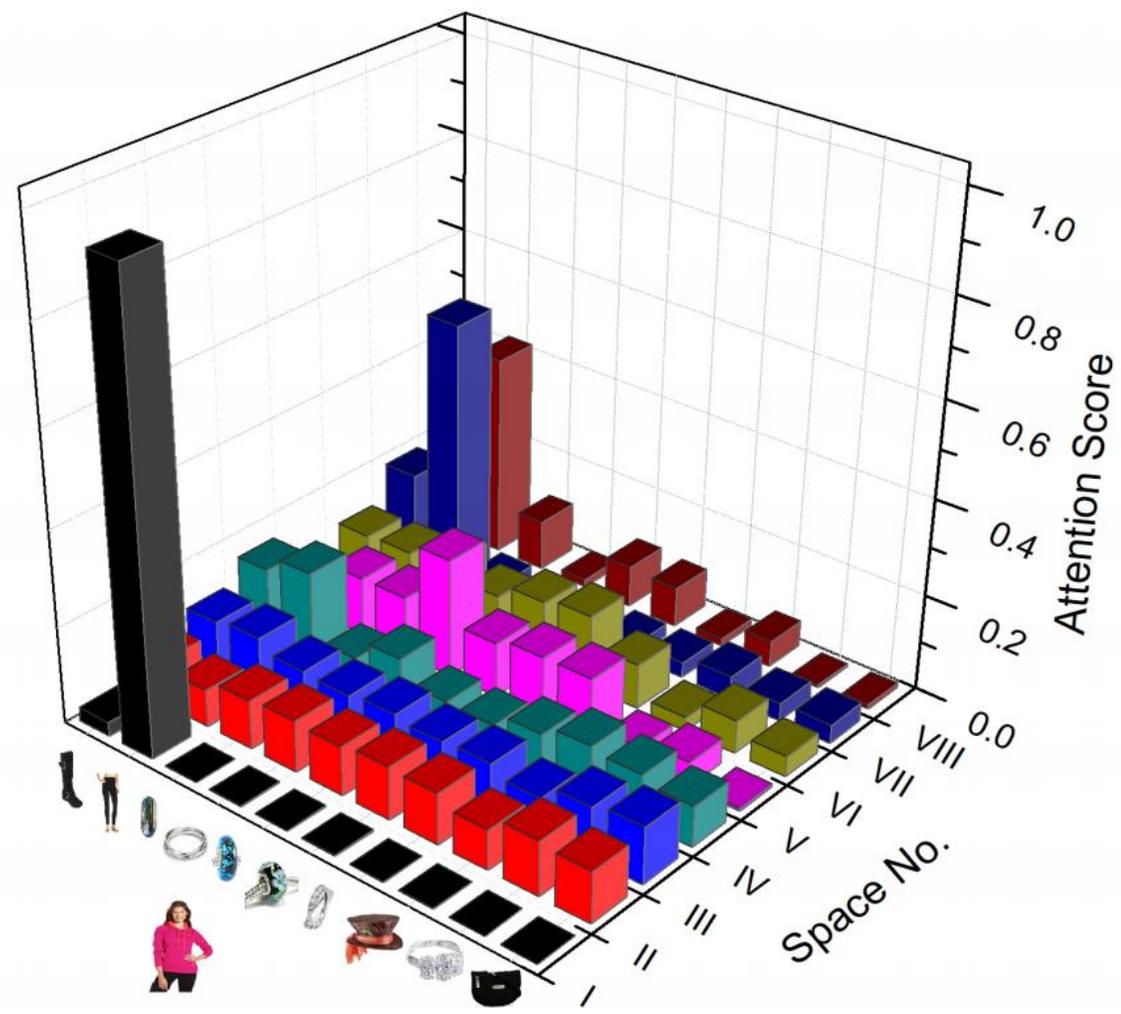
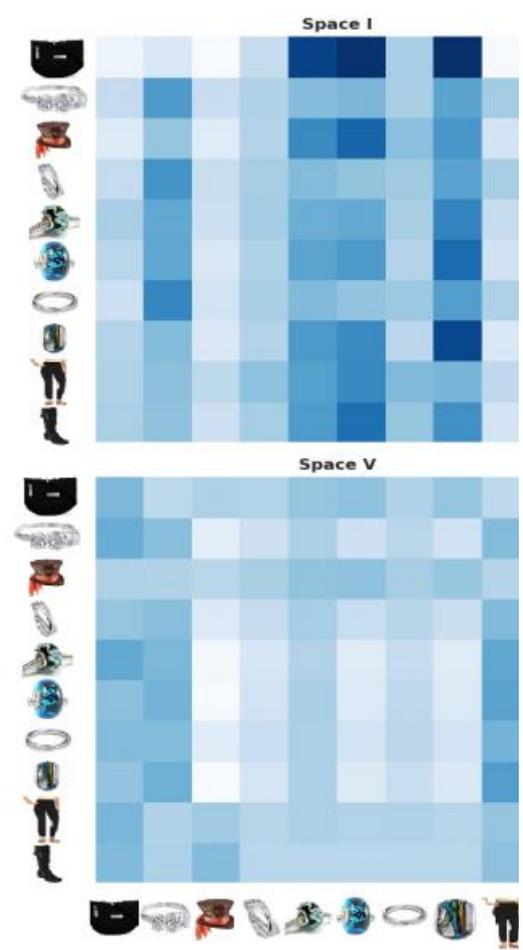
$$A_k = \text{softmax}(S_k W_k S^T)$$

$$C_k = A_k Q_k S$$

$$C = f(\text{concat}(C_1, C_2, \dots, C_k))$$



# 算法原理 ATRank case study



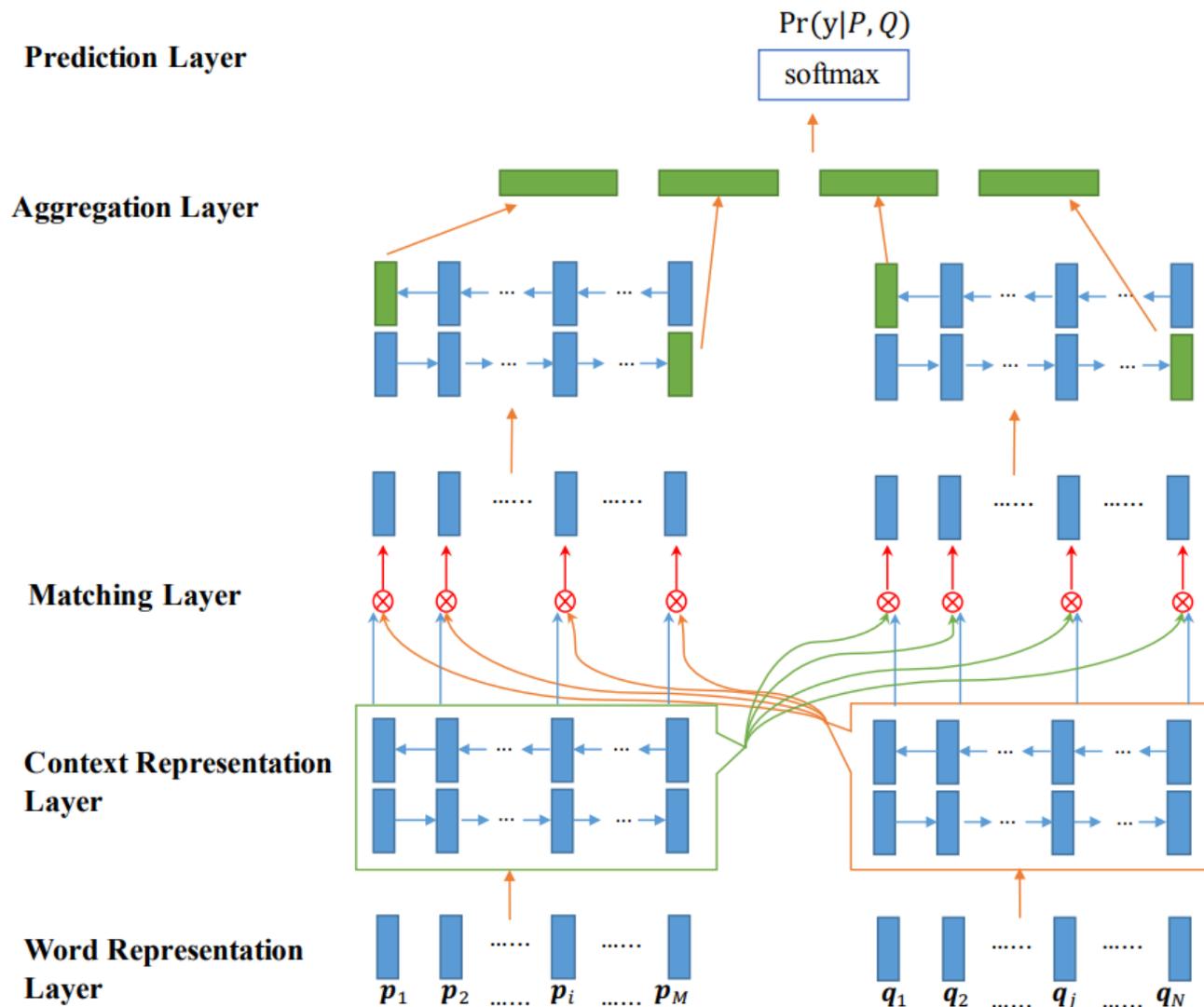


算法原理

T	利用双向多视角进行文本推理
I	文本短句
P	For { 1. 词、上下文嵌入表示 2. 多视角文本信息交互 3. 聚合与预测 }
O	文本推理结果

P	双向、多视角对句子进行建模
C	输入文本为短句
D	如何利用更充足的句子信息
L	IJCAI CCF A类

- 逻辑框架
  - 文本嵌入层
  - 匹配层
  - 聚合层
  - 输出层



- 双向多视角匹配层

- 双向匹配

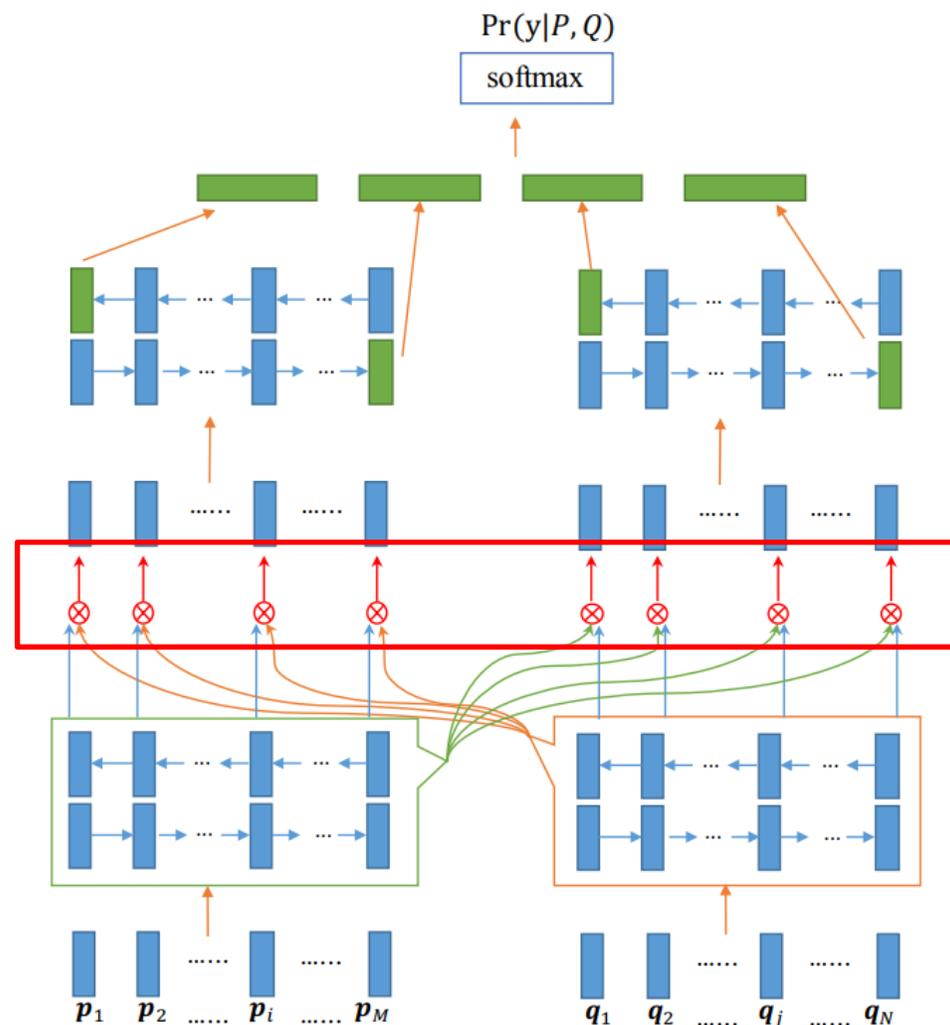
- $P \rightarrow Q$
- $Q \rightarrow P$

- 多视角匹配

$$m_k = \text{cosine}(W_k \circ v_1, W_k \circ v_2)$$

- 匹配策略

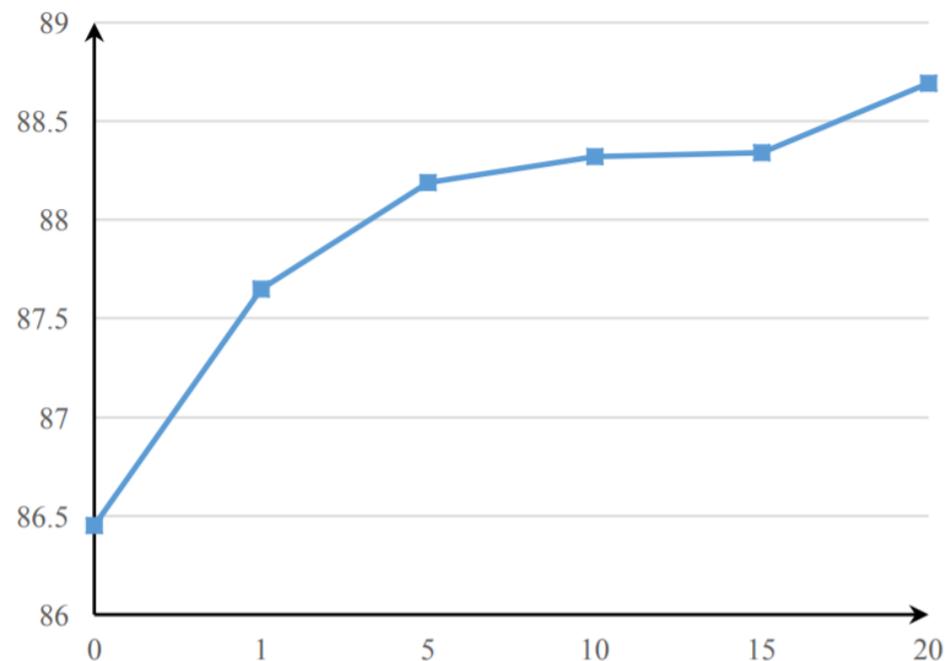
- Full Matching
- Maxpooling Matching
- Attentive Matching
- Max-Attentive Matching



## 实验结果对比

Models	Accuracy
[Bowman <i>et al.</i> , 2015]	77.6
[Vendrov <i>et al.</i> , 2015]	81.4
[Mou <i>et al.</i> , 2015]	82.1
[Rocktäschel <i>et al.</i> , 2015]	83.5
[Liu <i>et al.</i> , 2016b]	85.0
[Liu <i>et al.</i> , 2016a]	85.1
[Wang and Jiang, 2015]	86.1
[Cheng <i>et al.</i> , 2016]	86.3
[Parikh <i>et al.</i> , 2016]	86.8
[Munkhdalai and Yu, 2016]	87.3
[Sha <i>et al.</i> , 2016]	87.5
[Chen <i>et al.</i> , 2016] (Single)	87.7
[Chen <i>et al.</i> , 2016] (Ensemble)	88.3
Only $P \rightarrow Q$	85.6
Only $P \leftarrow Q$	86.3
BiMPM	86.9
BiMPM (Ensemble)	<b>88.8</b>

## 视角数量对结果的影响



## • 法律文本匹配应用

被告1因生意缺少资金，向原告借款，双方于当日签订了一张《借款合同》，担保人为被告2，担保方式为连带责任保证，约定借款金额60,000元，约定月利率2%，未约定借款期限。借款后被告1支付利息至2013年4月10日，此后经原告催讨未果，双方于2015年10月28日签订了还款承诺书，被告1结欠原告借款本金60,000元，利息38,400元，利息每月缴纳，原告未依约还款。

被告1以生意需要资金为由被告2担保，于2016年10月3日向原告借款11,500元，并于当日签订借条一张，双方约定月利率为2%，被告2自愿对上述债务承担连带清偿责任，借款后被告1未支付任何本息，后经过原告多次向被告1、被告2催讨本息，被告1、被告2拖延至今无果。被告1、被告2辩称，对原告主张的借款本金事实无异议，但利息是否有约定的事实，于庭审后7日内向法庭补充提交证据证实。

2015年3月1日，被告1在原告处借款80000元，约定月利率15%，被告2为该笔借款提供保证担保。后经原告多次催要未果，诉至法院，要求二被告共同偿还借款本金80000元，支付利息14400元，并负担本案诉讼费用。庭审中，原告增加诉讼请求，要求自2016年3月1日起按月利率15%计付逾期利息至借款本金清偿时止。原告提供如下证据支持其诉讼主张：借条一份。

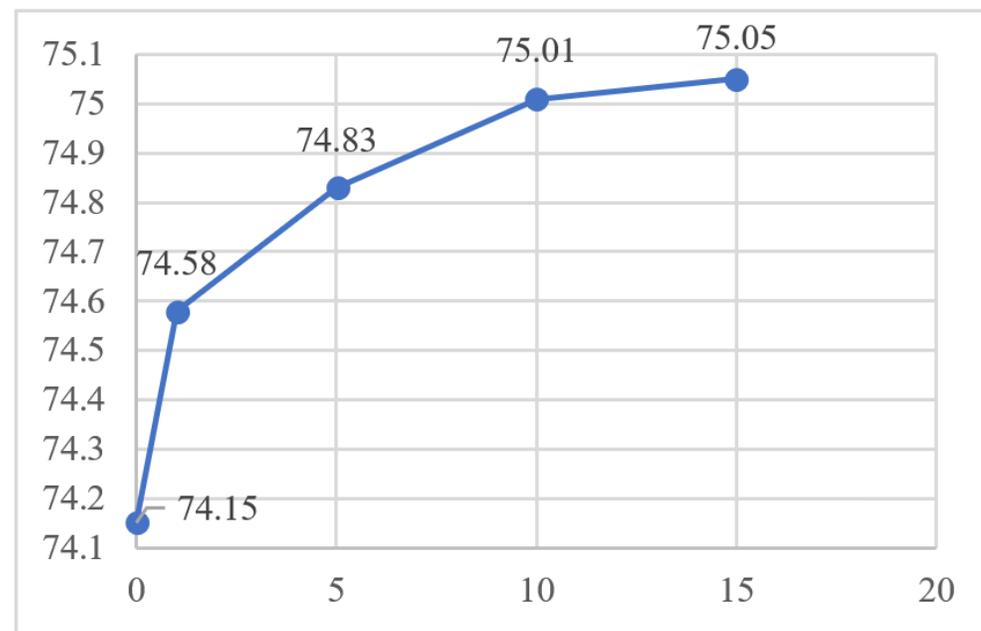
缓解单一语义空间下存在的语义中和现象

- 法律文本匹配应用

### 实验结果对比

Models	Valid Set Accuracy	Test Set Accuracy
CAIL2019-SCM 1st	70.07	72.66
LFESM	70.01	74.15
BERT-Siamese	67.73	71.81
w/o legal elements	70.33	74.46
w/o multi-perspective	70.43	74.67
<b>DLE-MP</b>	<b>70.67</b>	<b>75.05</b>

### 观点数量对结果的影响





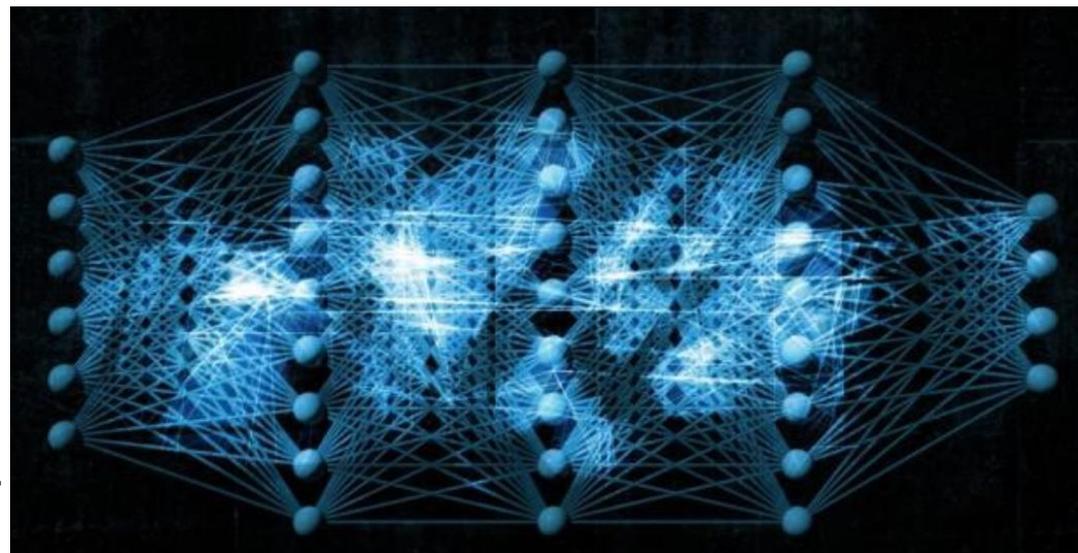
应用总结

- 优点

- 拆解多维度数据，处理异构信息
- 多视角间可并行计算

- 缺点

- 效率是多视角学习发展的最大瓶颈
  - 空间方面：多矩阵，多维度对内存要求较高
  - 时间方面：参数组合优化问题求解，时间复杂度高



- 算法的应用领域
  - 深度学习算法
    - 多源
    - 多特征子集
- 未来的发展
  - 时间与空间效率的提升
  - 可解释性

- [1] Xu C, Tao D, Xu C. A survey on multi-view learning[J]. arXiv preprint arXiv:1304.5634, 2013.
- [2] Wang Z, Hamza W, Florian R. Bilateral multi-perspective matching for natural language sentences[J]. arXiv preprint arXiv:1702.03814, 2017.
- [3] Zhou C, Bai J, Song J, et al. Atrank: An attention-based user behavior modeling framework for recommendation[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2018, 32(1).

大成若缺，其用不弊。  
大盈若冲，其用不穷。  
大直若屈。大巧若拙。  
大辩若讷。静胜躁，寒  
胜热。清静为天下正。

## 谢谢！

