

# 面向心血管健康管理领域的 知识图谱和知识库研究

汇报人： 邢俊劼 大三

指导老师： 朱其立 教授

上海交通大学电子信息与电气工程学院

# 目录

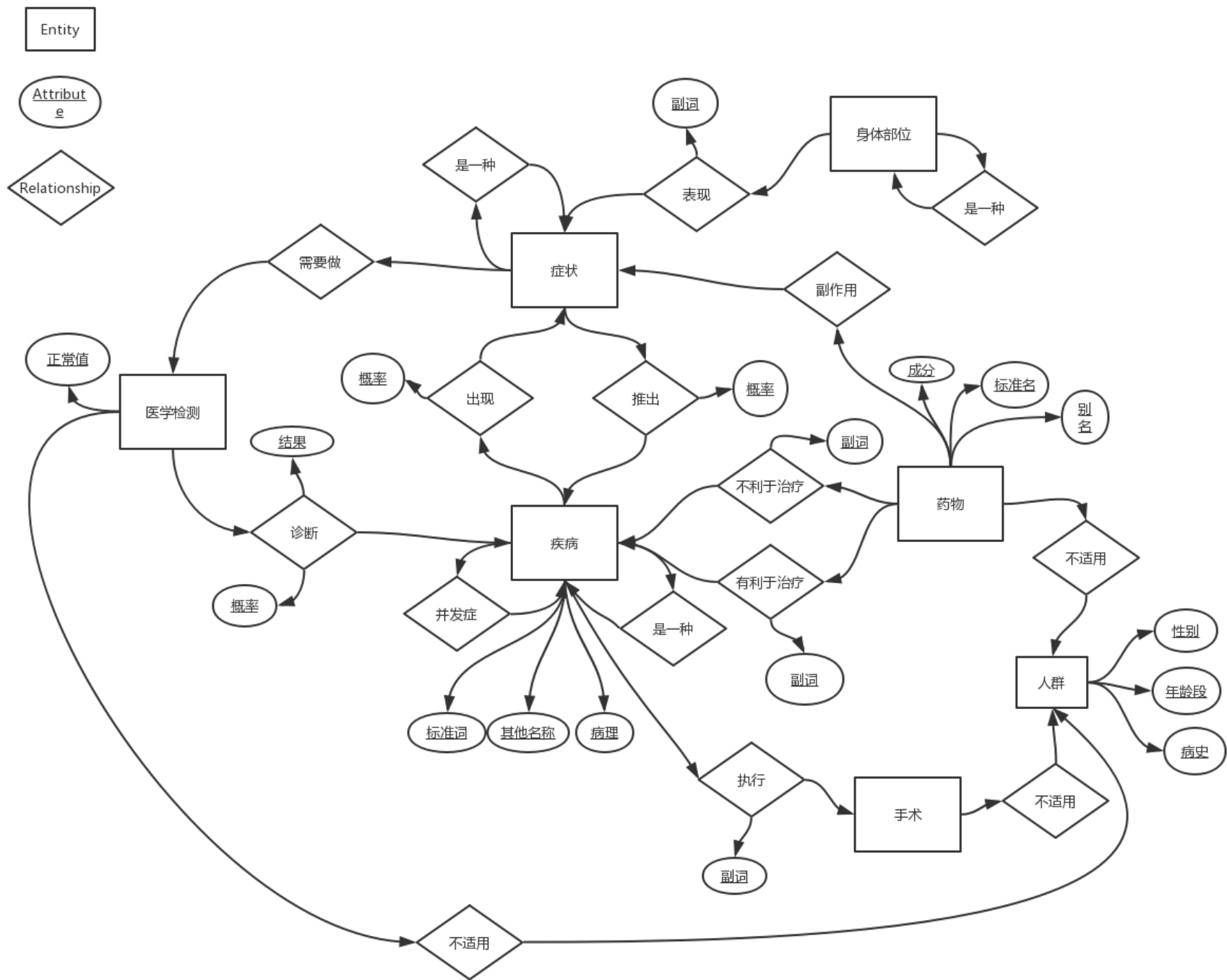
- 概述（列一下要做什么）
- 数据建模（讲解ER图，给出例子）
- 数据资源（列一下目前有的资源）
- 项目进展（目前使用的方法、结果、不足）
- 项目计划（未来的计划）
- 相关研究展示（梦雪）

# 概述

- 研究目标：
  - 心脑血管健康词典库和简单规则库
  - 举例：
    - 词典库： 高血压， 头晕。。。
    - 规则库： （头晕、血压高） => 高血压
- 研究语料来源：
  - 医学文献（论文、教材）
  - 社交媒体数据（寻医问药网等）

# 数据建模

- 建立实体——关系（ER）图展示数据模型
- 在对医疗信息领域进行了解之后，将数据实体定义为以下七类：
  - 疾病
  - 症状
  - 药物
  - 手术
  - 人群
  - 身体部位
  - 医学检测



# 数据建模

实体类型：疾病

疾病：高血压

标准词：高血压

其他名称：无

病理：遗传因素。。。

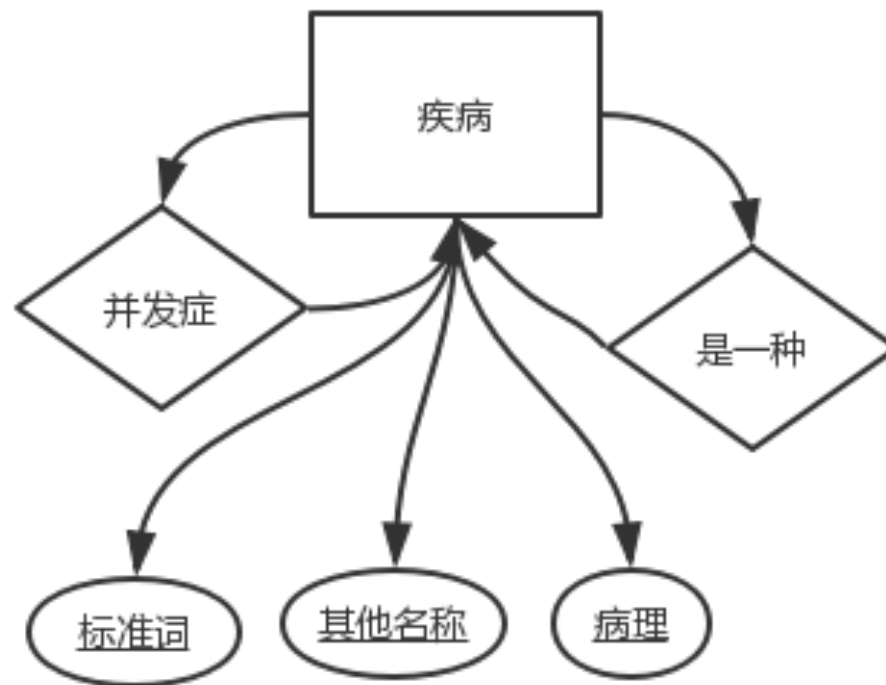
关系：

「冠心病」是「高血压」的（并发症）

「中风」是「高血压」的（并发症）

「原发性高血压」是一种「高血压」

「继发性高血压」是一种「高血压」



# 数据建模

实体类型：症状，身体部位

症状：瘙痒

身体部位：皮肤

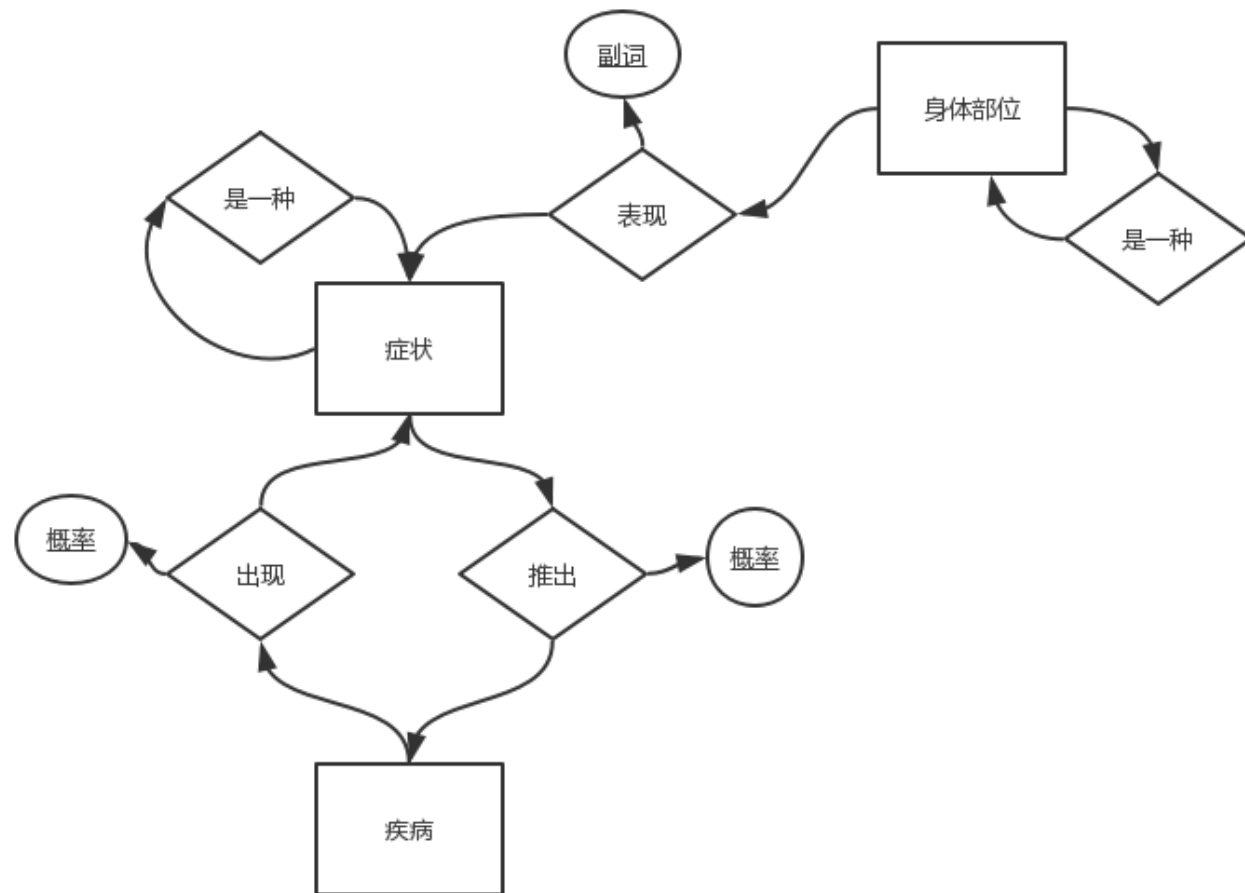
关系：

「皮肤」（身体部位）表现「瘙痒」（症状）

「头疼」（症状）是一种「疼痛」（症状）

「发热」（疾病）出现「头疼」（症状）

「皮肤瘙痒」（症状）推出「过敏性结膜炎」（疾病）



# 数据建模

实体类型：药物

别名：阿司匹林

标准词：乙酰水杨酸片

成分：乙酰水杨酸

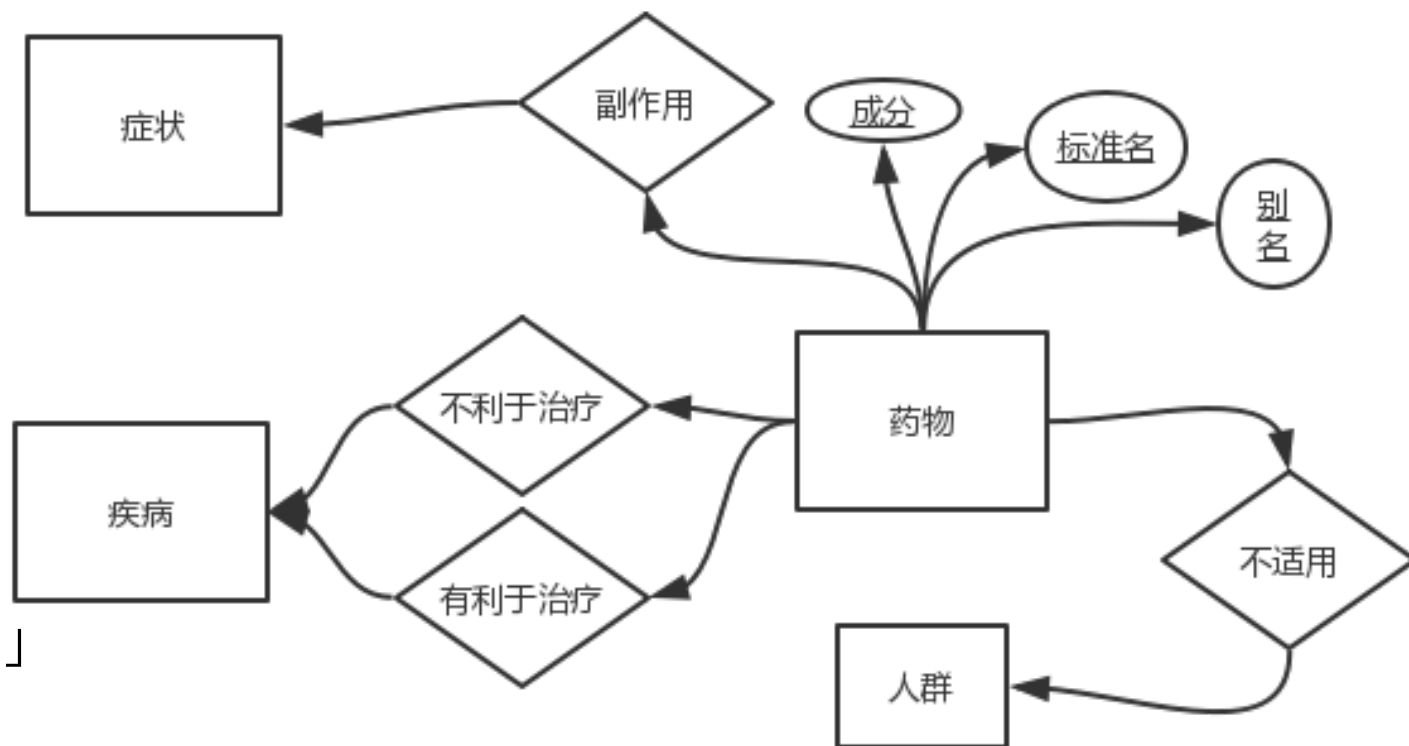
关系：

「阿司匹林」有利于治疗「发烧」

「阿司匹林」有利于治疗「风湿热」

「阿司匹林」副作用「消化道出血」（症状）

「阿司匹林」不适用于「水杨酸不耐受人群」





# 数据建模

实体类型：人群

人群：孕妇

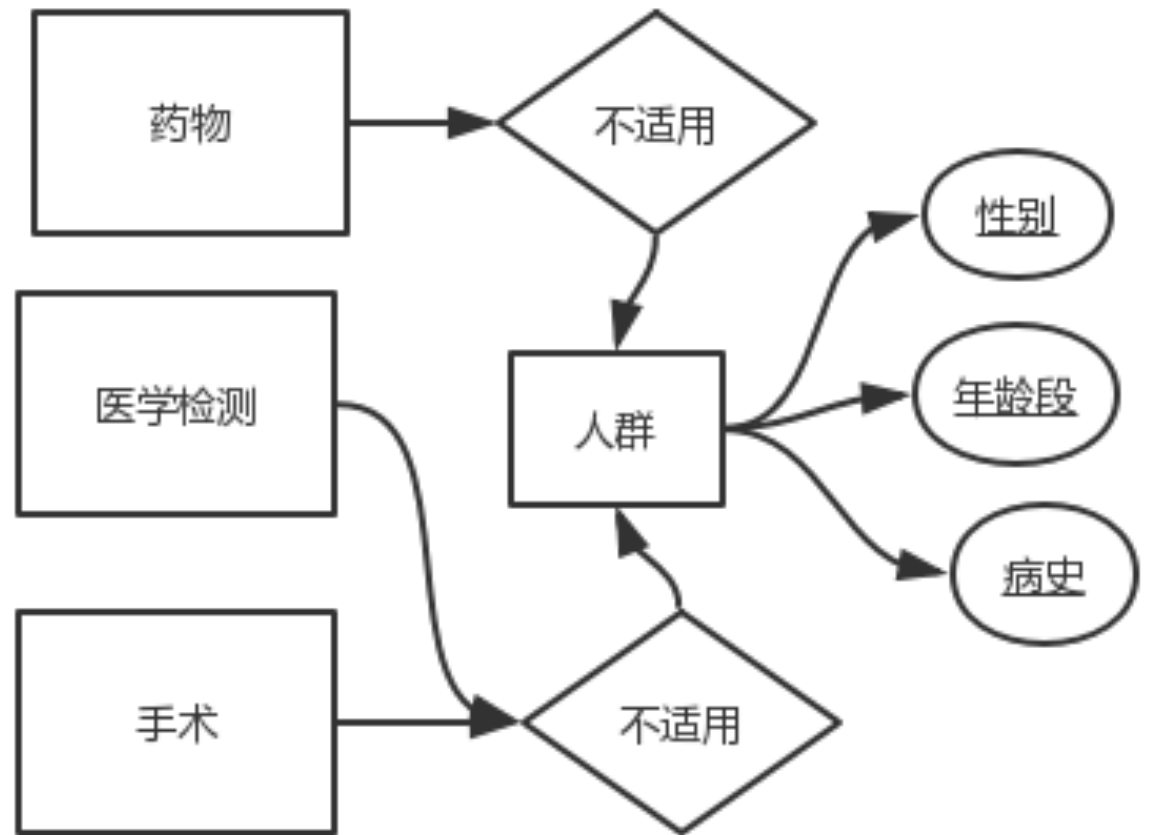
性别：女性

年龄段：孕期

关系：

「阿司匹林」不适用于「孕妇」

「孕妇」不适合「激光手术」



# 数据建模

实体类型：手术

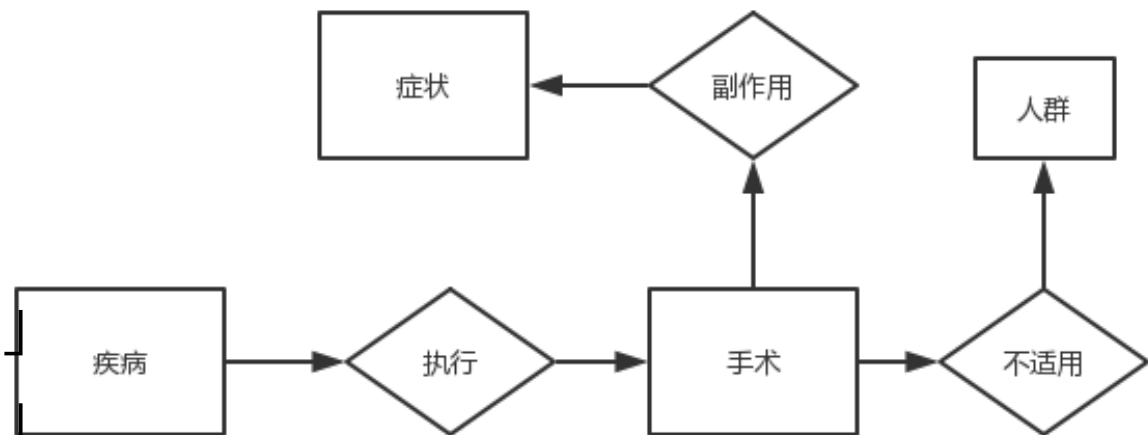
手术：激光近视手术

关系：

「近视」可以执行「激光近视手术」

「激光近视手术」不适用于「孕妇」

「激光近视手术」副作用「炫光症」



# 数据建模

实体类型：医学检测

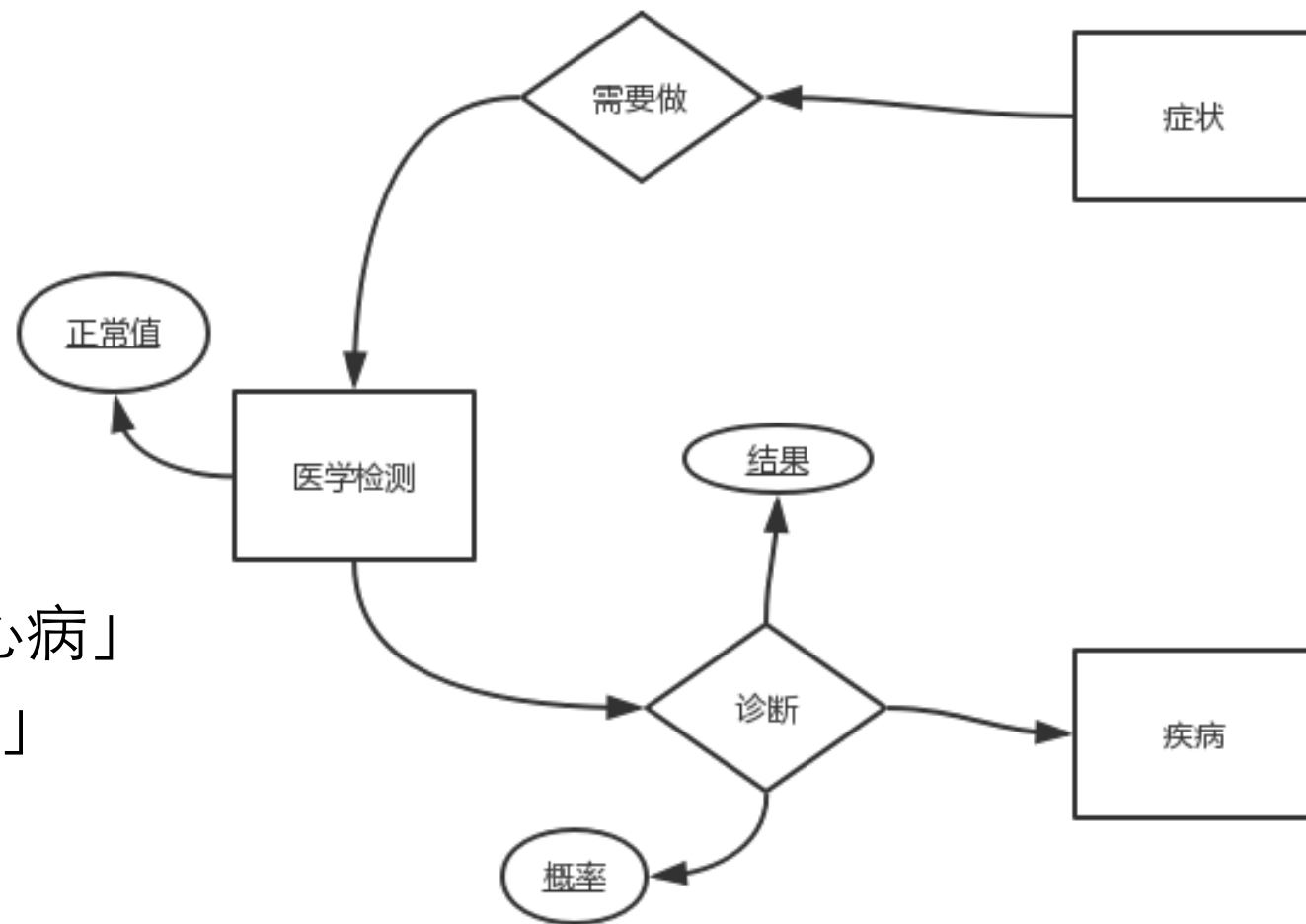
医学检测：心电图

关系：

「心悸」需要做「心电图」

「心电图结果」诊断「冠心病」

「孕妇」不适宜「过多B超」



# 数据资源

- 辞典库：
  - MedDRA: “MedDRA是一个内容丰富，详细的医学标准术语，促进人用医疗产品国际监管信息的共享。”（已使用）
  - SNOMED CT: 医学系统命名法 – 临床术语，是当前国际上广为使用的一种临床医学术语标准。（未使用）
- 医学文献：
  - 一共找到1500篇样本文献，后续我们还将爬取更多的医学文献数据
- 社交媒体：
  - 寻医问药网: 200GB问答数据，千万条

# 项目进展

- 辞典库建立

- MedDRA层次型数
- 优点：获取医学术
- 缺点：
  - 不够完善，缺乏口
  - 疾病、症状等为区

```
>>> get_one_path()
soc
|----code:10047065
|----name:血管与淋巴管类疾病
with 11 children

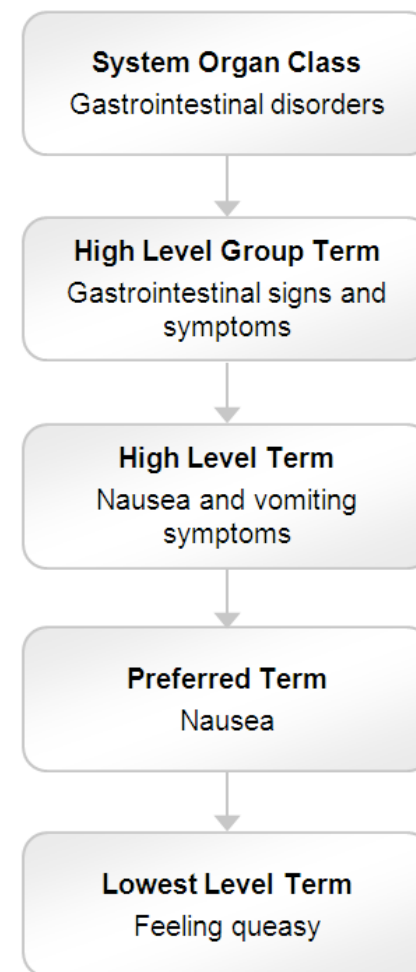
hlgt
|----code:10014523
|----name:栓塞及血栓
with 11 children

hlt
|----code:10038434
|----name:肾栓塞和血栓
with 5 children

pt
|----code:10038547
|----name:肾静脉栓塞
with 3 children

llt
|----code:10038547
|----name:肾静脉栓塞

>>>
```



# 项目进展

- 处理辞典：
  - 人工筛选
    - 工作量太大
  - 使用特征词将疾病区分出
    - '症', '不能', '综合征', '炎'
    - '硬化', '癌', '哮喘', '溃疡'
- 处理结果：
  - 保证疾病分类正确
  - 其余在排除手术、医学杂志
  - 疾病：21571项
  - 症状：27466项

0期甲状腺癌  
1型糖尿病  
1型糖尿病出现眼部并发症  
1.5型糖尿病  
18q综合征  
18q负综合征  
2型糖尿病  
2型糖尿病出现眼部并发症  
2型胰岛素依赖型糖尿病  
3型糖尿病  
3型糖尿病  
5q缺失骨髓增生异常综合征  
5q缺失骨髓增生异常综合征  
5q负综合征  
5q-综合征  
ABB1B2等维生素缺乏症  
A型胃炎  
A型血友病  
A型血友病  
A型血友病（因子VIII）  
A型血友病（因子VIII）  
A族β-溶血性产毒酿脓链球菌性蜂窝织炎  
A组β-溶血性产毒酿脓链球菌引起的中毒性休克综合征  
ACTH垂体瘤  
Aarskog-Scott综合征  
Aase综合征  
Ahumada-delCastillo综合征  
Aicardi综合征  
Alagille综合征  
Albright病  
Allgrove综合征  
Angelman综合征

Haglund畸形  
Henoch-Schonlein  
Holmes震颤  
Hutchinson夏令痒疹  
I型糖尿病伴其他昏迷  
I型糖尿病伴酮症酸中毒  
I型糖尿病控制不良  
I型糖尿病肾脏表现  
I型超敏反应  
I型超敏性  
I型超敏性（未特指）  
IA间隔缺损  
IC型支架移植植物内漏  
ICU获得性衰弱  
II型糖尿病伴其他昏迷  
II型糖尿病伴酮症酸中毒  
II型糖尿病控制不良  
II型糖尿病肾脏表现  
II型超敏反应  
II型超敏反应  
IIA型支架移植植物内漏  
IIB型支架移植植物内漏  
III型免疫复合物介导的反应  
III型超敏反应  
IIIA型支架移植植物内漏  
IIIB型支架移植植物内漏  
IV型反应  
IV型超敏反应  
IX因子抑制  
IX因子抑制复发  
IgA缺乏  
IgG亚型缺乏

# 项目进展——医学文献

## • 使用OCR工具处理文献

### 第一章 概述

蔡映云

随着科学技术的日益发展,生活水平的不断提高以及社会福利保障体制的完善,人类的平均寿命日趋延长。预测 2025 年全球将进入老年型社会。中华人民共和国国家统计局公布的“2010 年第六次全国人口普查主要数据公告(第一号)”的资料显示,2010 年全国总人口 13.7 亿,其中 60 岁及以上人口占 13.26%,65 岁及以上人口占 8.87%。与 2000 年第五次全国人口普查相比,60 岁及以上人口的比例上升 2.93 个百分点,65 岁及以上人口上升 1.91 个百分点。面临人口老龄化的现状,老年医学遇到前所未有的挑战和责任。

#### 【老年人的年龄划分】

从衰老的进程来看,老年期可分为三个阶段,即老年前期、老年期和长寿期。按联合国区域划分,亚太地区把 60 岁以上定为老年人。我国通用标准是 45~59 岁为老年前期(中年人);60~89 岁为老年期(老年人);其中 80 岁以上称高龄老人;≥90 岁为长寿期(长寿老人),其中≥100 岁以上称百岁老人。北美和多数欧洲国家则以 65 岁为界,超过此限者定为老年人。

在老年医学中,有三种表达年龄的方法,即时序年龄、生物学年龄和心理年龄。每一个个体都存在这三种年龄,从不同角度反映出一个人的生命状态。时序年龄又叫立法年龄,是出生以后所经历时期长短的个体年龄,时序年龄只能表示老年人的总体衰老程度和状态。生物学年龄则取决于人体结构和功能的衰老程度,表示人体组织结构和生理功能的实际衰老程度。生物学年龄从客观上反映了老年人的个体差异。心理年龄是通过测定个体的记忆力、心算能力及过去记忆力,将它与以前智能高峰期的结果比较,测出智能衰退率。一般 60 岁以上老人的衰退率约为 20%,70 岁的老人约为 30%。把心理学年龄与时序年龄对

照,就能看出智力减退的情况。心理年龄还受到社会影响,老人精神愉快、身体健康,虽时序年龄较大,但显得年轻。

#### 【老年学与老年医学】

老年学(gerontology)是研究老年人各方面问题的一门学科,主要包括老年生物学、老年医学、老年社会学、老年心理学等。

老年医学(geriatric medicine)研究机体衰老的机制及其预防,老年人疾病的发病机制、预防、诊断和治疗以及老年人的卫生问题如心理、营养、康复等。老年医学是现代医学中一个重要的分支科学,又是老年学的重要组成部分。

老年医学包括老年基础医学、老年临床医学、老年流行病学、老年康复医学等学科。实际上各学科研究的内容是相同的,只是出发点或角度不同,或采用的方法和手段不同。而且各学科必须交汇融和、密切配合、协同作战,才能取得成绩。

(一) 老年基础医学 研究衰老的机制,探索延缓衰老的方法以及老年期疾病的病因和发病机制。使人类延缓衰老,延长健康生命。

(二) 老年临床医学 由于老年人在生理和病理上具有特殊性,在诊治中与青年人存在较大差别。给老年人疾病的诊断和治疗带来困难,治疗的目标、预后和转归也与中年人不同。老年临床医学就是研究对于老年患者进行正确诊断和合理治疗的科学。

(三) 老年流行病学 探索影响人体衰老的因素和老年常见病的分布状况和发病规律,从而寻求延缓衰老、预防疾病、提高全社会老年人生命质量的对策。并进一步实施干预措施,评估干预的效果。与老年流行病学密切相关的

9

#### 第九篇老年医学

##### 第一章概述

蔡映云

随着科学技术的日益发展,生活水平的不断提窃以及 社会福利保障体制的完备,人类的平均寿命H趋延长。预 测2025年全球将进入老年型社会。中华人民共和国国家 统计局公布的“2010年第六次全国人口普查主要数据公告 (第一号)”的资料显示,2016 年全国总人口 13.7 亿,其中 60 岁及以上人口占13.26%,65 岁及以上人口占8.87%。与2000年第五次全国人口皆费相比,60 岁及以上人口的 比例上升2.93个百分点,65 岁及以上人口上升1.91个百 分点。面临人口老龄化的现状,老年医学遇

##### 【老年人的年龄划分】

照,就能看出智力减退的情况。心理年龄还受到社会影 响,老人精神愉快、身体健康,虽时序年龄较大,但显得 年轻。

##### 【老年学与老年医学】

老年学(gerontology)是研究老年人各方面问题的一门 学科,主要包括老年生物学、老年医学、老年社会学、老年心 理学等。

老年医学 (geriatric medicine)研究机体衰老的机制及 其预防,老年人疾病的发病机制、预防、诊断和治疗以及 老年人的卫生问题如心理、营养、康复等。老年底学是现 代医学中一个重要的分支科学,又是老年学的重要组成部分

从衰老的进程来看,老年期可分为5个阶段,即老年前 期、老年期和长寿期。按联合国区域划分,亚太地区把60 岁以上定为老 人。我国通用标准是45~59岁为老年前 期(中年人);60~59岁为老年期(老年人); 其中80岁以上 称尚龄老人090岁为长寿 期(长寿老人), 其中>100岁以 上称百岁老人。北美和多数欧洲国家则以65岁为界,超过 此限者定为老年人。

在老年医学中,•有三种表达年龄的方法,即时序年龄、 生物学年龄和心理年龄。每一个个体都存在这三种年 龄,从不同角度反映出一个人的生命状态。时序年龄又叫 立法年龄,是出生以后所经历时期长短的个体年龄,时序年 龄只能表示老年人的总体 老程度和状态。生物学年龄则 取决于人体结构和功能的衰老程度,表示人休组织结构和 生理功能的实际袋老程度。生物学年龄/ 客观上反映r老 年人的个体差异。心理年龄泉通过测定个体的牢记忆力、 心符•能力及过去记忆力,将它4以前智能高峰期的结果 比较,测出智能衰退率。一般60岁以[•。老人的衰退率约为 20%,70岁的老人约为30%。把心理学年龄与时序年龄对 部分。

老年医学包括老年基础医学、老年临床医学、老年流行 病学、老年康复医学等学科。实际上各学科研究的内容是 相同的,只是出 发点或角度不同,或采用的方法和手段不 同。而且各学科必须交汇融和、密切配合、协同作战•才能 取得成绩。

(一)老年基础医学研究衰老的机制,探索延缓衰老 的方法以及老年期疾病的病因和发病机制。使人类延缓衰 老,延长健康的生命。

(二) 老年临床医学由于老年人在生理和病理上具 荷特殊性,在诊治中与中青年存在较大差别。给老年人 疾病的诊断和 治疗带来困难,治疗的目标、预后和转归也与 中年人不同。老年临床医学就是研究对于老年患#进行正 确诊断和合理治疗的科 学。

(三) 老年流行病学探索影响人体衰老的因素和老 年常见病的分布状况和发病规律•从而寻求延缓衰老、预防 疾病、提商全 社会老年人生命质位的对策。并进••步实施

丨•预措施,评估干预的效聚。与老年流行病学密切相关的

215

#### 第九篇老年医学-----

学科主要有:临床|装学、基础医学、统计学和概率论、社会学 和心理学。

(四)老年预防医学研究老年人常见病的病因、危险 因素和保护因素,探讨有效的预防措施,如卫生宣教、合理 的生活方式和营 养饮食。不仅预防病残,还要提高生活质 1:,使老人为社会继续贡献。预防医学内容包括老年流行 病学、营养学、运动医学、 养生保健、心理卫生及健康教 育等。

# 项目进展——医学文献

- OCR识别软件分析
  - 现阶段的OCR软件已经能够比较好的识别图片并将其转化为文字 可以直接使用
  - 其缺点就是识别速度较慢 需要人工导入文献
- 目前所遇到的问题
  - 文本内容很长 导致疾病和症状隔离了多个句子之间
  - 难以找到句子和文本之间的真实关系
  - 症状和疾病的内容相对稀疏
  - 跟疾病有关的医学文献需要进行分词处理 这是特殊领域的分词 现有的分词软件效果不佳



# 项目进展——社交媒体

- 匹配方法：
  - 根据疾病辞典和症状辞典进行完全匹配
  - 错别字检查（根据拼音）
    - 例如： 头晕木眩 => 头晕目眩
- 基于社交媒体的关系抽取：
  - 根据观察，绝大多数回答都是针对问题，因此认为问题中的“症状”和回答中的“疾病”具有先天性关联
  - 大数据分析下，噪声将消失

# 项目进展——社交媒体

```
q:
    你好，我想问我现在得耳鸣耳聋头晕目眩一个多月了。
find:
    ['耳鸣', '耳聋', '头晕目眩']
a:
    问题分析：你好，根据你所说的症状，你的这种情况考虑是美尼尔综合症。意见建议：此病是由于内耳迷路水肿引起的。治疗上建议你用培他定和西比灵甘露醇，祝你健康。
find:
    ['水肿']
美尼尔综合症
```

## 美尼尔氏综合症- 维基百科，自由的百科全书

<https://zh.wikipedia.org/zh-hans/美尼尔氏综合症> ▼

**美尼尔氏综合症**（Ménière综合症），也稱為梅尼爾氏症、梅尼埃氏症、或俗稱的耳水不平衡。该病是由法国医生柏斯貝·美尼尔（Prosper Ménière）於1861年首先报告的。

診斷 · 症状 · 病理 · 治療

```
>>> trace_root_name('美尼尔氏综合症')
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
  File "/Users/xingjuniie/Documents/adapt/health/code/build_lexicon.py", line 188, in trace_root_name
    raise Exception("fail to locate {} in llt_list".format(name))
Exception: fail to locate 美尼尔氏综合症 in llt_list
>>> trace_root_name('梅尼埃病')
梅尼埃病-->梅尼埃病-->各种眩晕（不另分类）-->神经类疾病（不另分类）-->各类神经系统疾病
>>>
```

# 项目进展——社交媒体

- 上述所反映的问题：
  - 疾病的多种表述方式
  - 医生、病人都会表述出错
  - 辞典库不完善
- 解决方案：
  - 多种同义表达的映射
  - 匹配时容错

# 项目进展——社交媒体

- 在寻医问药网一年的问答数据（56GB）进行匹配分析之后，截取以下结果

老年性瘙痒症

带状疱疹, 不舒服, 发红, 干燥, 尿糖, 皮疹, 乏力, 胃疼, 疹, 流血, 痂, 高血压, 腹泻, 过敏, 瘙痒, 癣, 发痒, 发痒的, 疱疹, 热, 发作, 全身瘙痒, 结痂

咽鼓管疾病

听力下降, 头晕, 感冒, 耳聋, 耳鸣, 口干, 震动

# 项目进展——社交媒体

## 上呼吸道感染

咳嗽, 40709, 0.2, 0.07; 感冒, 21619, 0.11, 0.06; 痰, 19434, 0.1, 0.06; 发烧, 17143, 0.08, 0.05; 流鼻涕, 10077, 0.05, 0.08; 鼻涕, 6182, 0.03, 0.05

## 宫颈炎

白带, 26827, 0.27, 0.06; 出血, 9878, 0.1, 0.03; 宫颈糜烂, 9514, 0.1, 0.1; 疼痛, 3914, 0.04, 0.01; 流血, 2780, 0.03, 0.03; 外阴, 2706, 0.03, 0.02

## 糖尿病

糖尿, 125465, 0.39, 0.66; 高血压, 8498, 0.03, 0.12; 糖尿病的, 7303, 0.02, 0.72; 头晕, 6798, 0.02, 0.03; 口干, 5460, 0.02, 0.15; 口渴, 5449, 0.02, 0.4

## 呼吸道感染

咳嗽, 26899, 0.22, 0.05; 感冒, 12818, 0.1, 0.04; 痰, 12668, 0.1, 0.04; 发烧, 11293, 0.09, 0.03; 流鼻涕, 4682, 0.04, 0.04; 鼻涕, 2952, 0.02, 0.02

# 项目计划

- 使用更多的NLP方法扩充辞典库，而不是依托于现有的医学词典
  - 建立基于医学语料（论文、社交媒体）的中分分词工具
    - 现有的中文分词工具，Stanford NLP等在项目语料上表现不足
  - 建立基于医学语料的NER命名实体识别工具
    - 通过NER扩充辞典库，扩充口语化描述
- 使用跨文本关系提取方法
  - 现在的关系抽取建立在QA系统的特殊性上，不可靠

# 项目计划——Reference

- Zheng, Suncong, et al. "Joint Extraction of Entities and Relations Based on a Novel Tagging Scheme." *arXiv preprint arXiv:1706.05075* (2017).
  - 提供了一个同时抽取实体和关系的方法
- Peng, Nanyun, et al. "Cross-Sentence N-ary Relation Extraction with Graph LSTMs." *Transactions of the Association for Computational Linguistics* 5 (2017): 101-115.
  - 提供了一个跨文本的关系抽取方法（可能适用于QA中的关系抽取）

# 相关项目——药物的副作用研究

- Motivation:
  - 药品说明书上所写的药品副作用往往不全面，且研究药品副作用的价格昂贵，耗时长。
  - 因此，我们考虑在社交网络上提取有用的药品副作用信息。
  - 例如：倍他乐克，



# 相关项目——药物的副作用研究

- 数据来源
  - 社交网络：寻医问药+好大夫
- 第一步需要对所有帖子进行预处理：
  - 保证数据集中的每句话包含至少一种疾病症状和至少一种药物

# 相关项目——药物的副作用研究

- 定义为 二分类 问题
- positive:
  - 句子中的疾病是由药物导致的
  - 如：吃了倍他乐克(drug)后感到头晕(ADR)。
- negative:
  - 句子中的药物是用来治疗疾病的
  - 如：去医院查出心律失常(ADR)，医生开了倍他乐克(drug)。

# 相关项目——药物的副作用研究

- ADR词典建立

- 基础词典库：根据MedDra建立起基础ADR(Adverse Drug Reaction)词典库
- 延伸词典库：用副词对上述词典库进行扩建：
- 头晕 -> 头有点晕
- 头很晕
- .....

# 相关项目——药物的副作用研究

- 寻找副词

1. Use the regular expression to get the “adv”s:

Ex: ‘头(.{1,5})晕’

We could get: 有点，很，非常.....

- 使用正则匹配获取语料库中出现的副词

# 相关项目——药物的副作用研究

- 药品词典库建立
  - 选定与治疗心脑血管疾病有关的药品共多少种
  - 比如“倍他乐克”
  - 对其进行同音同义字的扩展：

贝塔乐可

北塔乐客

倍它乐客

.....

# 相关项目——药物的副作用研究

SVM中所使用的七个特征值：

- Verbs before the drugs, e.g. “服用(take)” in “服用倍他乐克(take Betaloc)”；
- Verbs before the conditions, e.g. “感到(feel)” in “感到头晕(feel dizzy)”；
- Verbs after the conditions, e.g. “好转(improved)” in “头疼好转(headache improved)”；
- Preposition, conjunction and noun of locality, e.g. “因为(because of)” in “因为头疼(because of headaches)” and “后(after)” in “服用倍他乐克后(after taking Betaloc)”；
- Punctuations that surround drugs and conditions;
- The number of other drugs and other conditions between the drug and condition of interest;
- A Boolean value that indicates whether condition appears in front of the drug or not.

# 相关项目——药物的副作用研究

训练过程（运用bootstrapping）：

- 一开始手动标的一小部分数据集S1
  - 用S1训练出一个SVM分类器，记为M1
- 让M1对所有句子进行分类。如果M1说此句话为pos，且根据此药品说明书，表明疾病确实是此药品的副作用。则把这句话加入到positive training set；如果M1说此句话为neg，且根据此药品说明书，表明药品是治疗该疾病的。则把这句话加入到negative training set；其他的情况，则不做任何操作
- 用新的training set训练出一个新的SVM分类器M2.
- 重复上述的迭代操作，直到数据集收敛。

# 相关项目——药物的副作用研究

HMM baseline

- Use the method in the paper of Sampathkumar, H. et al.
- 5 States:  
0: drug   1: ADR   2: indication   3: keyword   4: other
- Keywords are those words which denote the causal relationship of a drug causing a side-effect like “引起” , “导致” , “感到” .



# 相关项目——药物的副作用研究

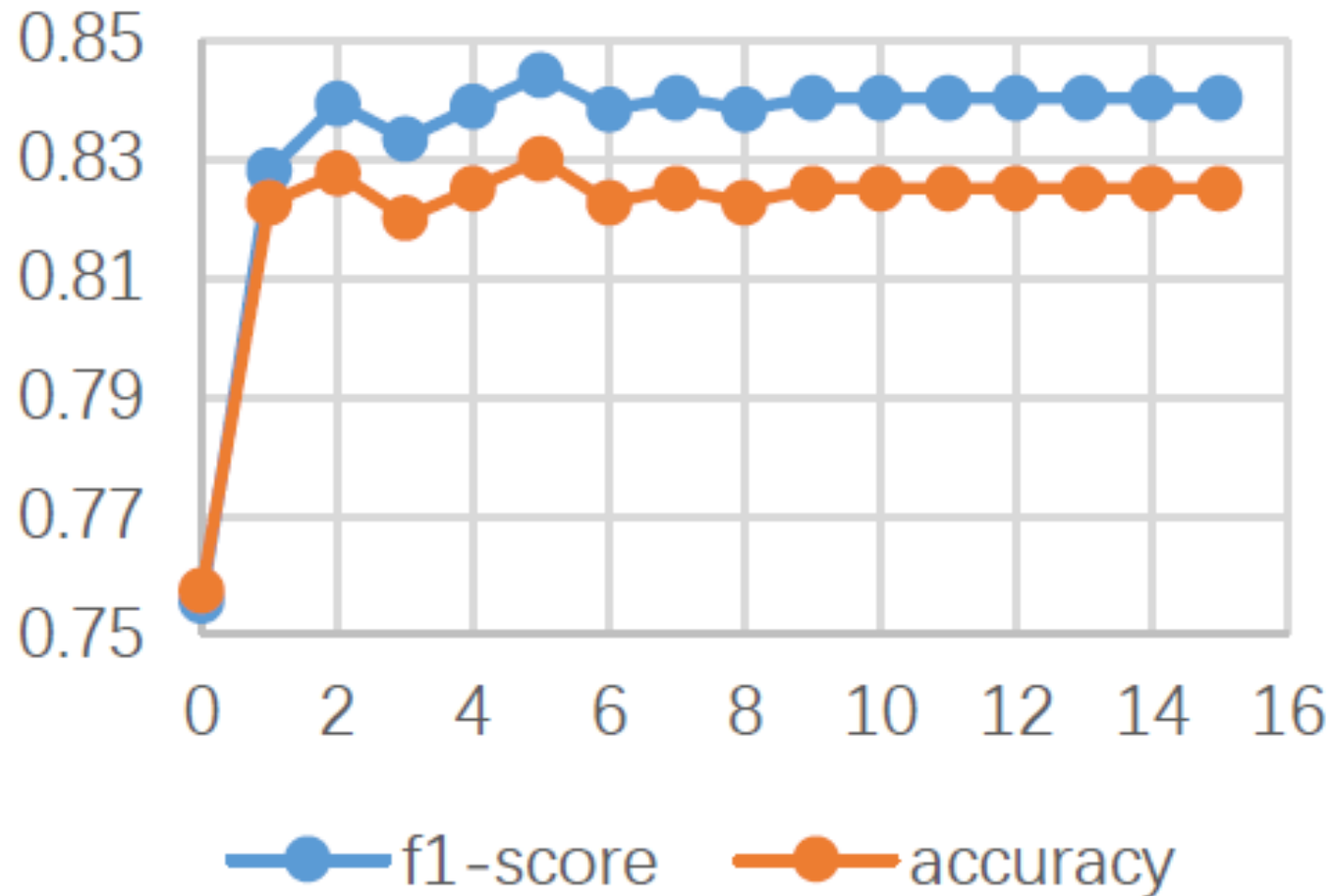
CRF baseline

- Use the method in the paper of Nikfarjam, A. et al.

According to the original paper, the features are as followed:

- **Context features** (current token, three preceding tokens and three following tokens)
  - **ADR lexicon**: a binary feature that shows whether or not the current token exists in the ADR lexicon.
  - **POS**: part of speech of the token, which was generated using Stanford Parser
  - **Negation**: This feature indicates whether or not the token is negated.

F1-score and accuracy of the new SVM classifier at each iteration



# 不同分类器的效果

- 结果展示:

Methods	positive pairs	negative pairs	recall	precision	F1-score
Manual labels (Pattern-based)	24/100	97/100	0.24	0.889	0.378
Semi-supervised (Pattern-based)	76/100	89/100	0.76	0.874	0.813
Manual labels (HMM)	62/100	85/100	0.62	0.805	0.700
Semi-supervised labels (HMM)	87/100	54/100	0.87	0.654	0.747
Manual labels (CRF)	86/100	75/100	0.86	0.775	0.815
Semi-supervised labels (CRF)	98/100	34/100	0.98	0.598	0.742
Manual labels (SVM)	68/100	87/100	0.68	0.840	0.752
Auto labels from inserts (SVM)	86/100	58/100	0.81	0.698	0.75
Semi-supervised labels (SVM)	86/100	79/100	0.86	0.804	<b>0.831</b>

# SVM各feature的重要性

SVM Features	positive pairs	negative pairs	R	P	F1	accuracy
All	184/200	148/200	0.92	0.74	<b>0.820</b>	<b>0.83</b>
All – verbs before drugs (feature 1)	175/200	152/200	0.875	0.76	0.813	0.818
All – verbs before conditions (feature 2)	184/200	147/200	0.92	0.735	0.817	0.828
All – verbs after conditions (feature 3)	175/200	<b>153/200</b>	0.875	<b>0.765</b>	0.816	0.82
All – preposition, conjunction and noun of locality (feature 4)	<b>187/200</b>	144/200	<b>0.935</b>	0.72	0.814	0.828
All – punctuations (feature 5)	169/200	131/200	0.845	0.655	0.738*	0.75
All – number of other drugs and other conditions (feature 6)	180/200	141/200	0.9	0.705	0.791*	0.803
All – Boolean value (feature 7)	173/200	146/200	0.865	0.73	0.792*	0.798

# 模型所输出的10个发现的ADR

(Drugs) 药物	倍他乐克 (Betaloc)	耐信 (Nexium)	拜唐苹 (Glucobay)	氨茶碱 (Aminophylline)
副作用 (ADRs)	咳嗽(2.45%) (Cough)	咳嗽(1.77%) (Cough)	不适(3.31%) (Discomfort)	咳嗽(51.39%) (Cough)
	紧张(2.06%) (Nervous)	头晕(1.09%) (Dizziness)	无力(2.18%) (Acratia)	头晕(0.69%) (Dizziness)
	不适(4.04%) (Discomfort)	不适(2.30%) (Discomfort)	发热(1.48%) (Fever)	恶心(0.57%) (Nausea)
	心悸(2.82%) (Palpitation)	紧张(0.32%) (Nervous)	头晕(2.70%) (Dizziness)	心悸(0.26%) (Palpitation)
	头晕(5.52%) (Dizziness)	便秘(0.85%) (Constipation)	乏力(1.31%) (Weak)	呕吐(1.13%) (Emesis)
	疲劳(0.67%) (Fatigue)	疲劳(0.16%) (Fatigue)	瘙痒(0.87%) (Itching)	心动过速(0.19%) (Tachycardia)
	头痛(1.32%) (Headache)	失眠(0.50%) (Insomnia)	腹泻(1.13%) (Diarrhea)	心律失常(0.26%) (Arrhythmia)
	恶心(0.89%) (Nausea)	头痛(0.36%) (Headache)	低血糖(3.14%) (Hypoglycemia)	打鼾(0.22%) (Snore)
	便秘(0.16%) (Constipation)	心悸(0.11%) (Palpitation)	虚弱(0.52%) (Asthenia)	抽搐(0.22%) (Tic)
	瘙痒(0.14%) (Itching)	皮肤过敏(0.12%) (Skin allergy)	咳嗽(0.61%) (Cough)	紧张(0.12%) (Nervous)

谢谢！