

附属说明材料:

1、往年的 AGAC 比赛网站: <https://sites.google.com/view/bionlp-ost19-agac-track>
比赛结果:

Task 1

	参赛者	Precision	Recall	F-score	主要的 NLP 方法
1st	参赛者 1	0.63	0.56	0.60	Bert, joint learning
*	Baseline	0.50	0.51	0.50	Bert, joint learning
2nd	参赛者 2	0.36	0.59	0.45	Bert, CNN, Bi-LSTM
3rd	参赛者 3	0.55	0.28	0.37	CRF, Bi-LSTM

上表显示的是任务一最终的比赛结果, 第一组参赛者使用的是“JFB-NER”模型, 这是一个参数微调 bioBert 的联合学习模型。第二组参赛者使用了一个分层多任务学习模型来进行命名实体识别和关系提取。第三组参赛者使用了具有语言特征的 Bi-LSTM CRF, 并在 3 个数据拆分上集成 3 个最佳模型。

Task 2

	参赛者	Precision	Recall	F-score	主要的 NLP 方法
1st	参赛者 1	0.40	0.31	0.35	Bert, CNN, Bi-LSTM
2nd	参赛者 2	0.61	0.16	0.25	Bert, joint learning
3rd	参赛者 3	0.05	0.02	0.03	SVM

任务二的比赛结果如上表, 第一组参赛者使用了一个分层多任务学习模型, 在关系提取部分, 该模型与任务一命名实体识别部分具有相同的编码层。第二组参赛者使用了一个简单的微调 bioBert, 称为“SB-RE”。第三组参赛者将任务二转换为分类模型, 使用传统的支持向量机, 得到的 F 值为 0.03。

Task 3

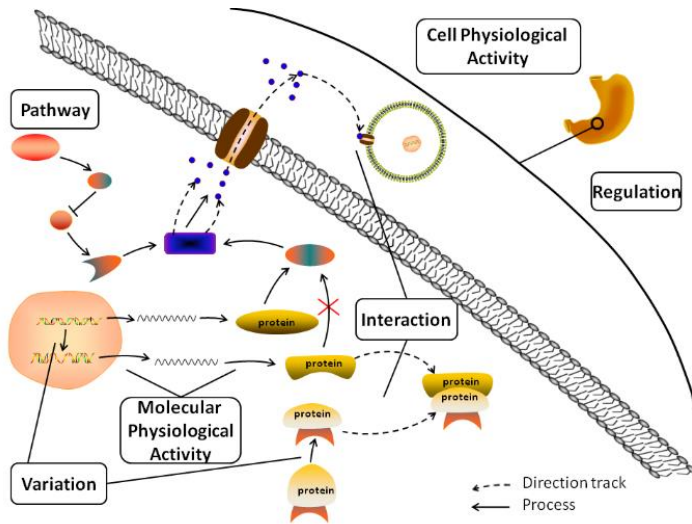
	参赛者	Precision	Recall	F-score	主要的 NLP 方法
*	Baseline	0.72	0.59	0.65	Bert, joint learning
L	参赛者 1	0.26	0.20	0.23	Bert

在任务三中有只有一组参赛者, 使用的是 Bert 模型提取三元组, 将提及到的文本及其文本上下文编码为两个连续的序列, 然后使用单个线性层将它们的关系分为五个类别。

更多比赛结果见已发表文献:

Yuxing Wang, Kaiyin Zhou, Mina Gachloo, Jingbo Xia*. An Overview of the Active Gene Annotation Corpus and the BioNLP OST 2019 AGAC Track Tasks. BioNLP Open Shared Task 2019, workshop in EMNLP-IJCNLP 2019. Page: 62-71, Hong Kong, 2019

2、AGAC 语料库标注指南说明

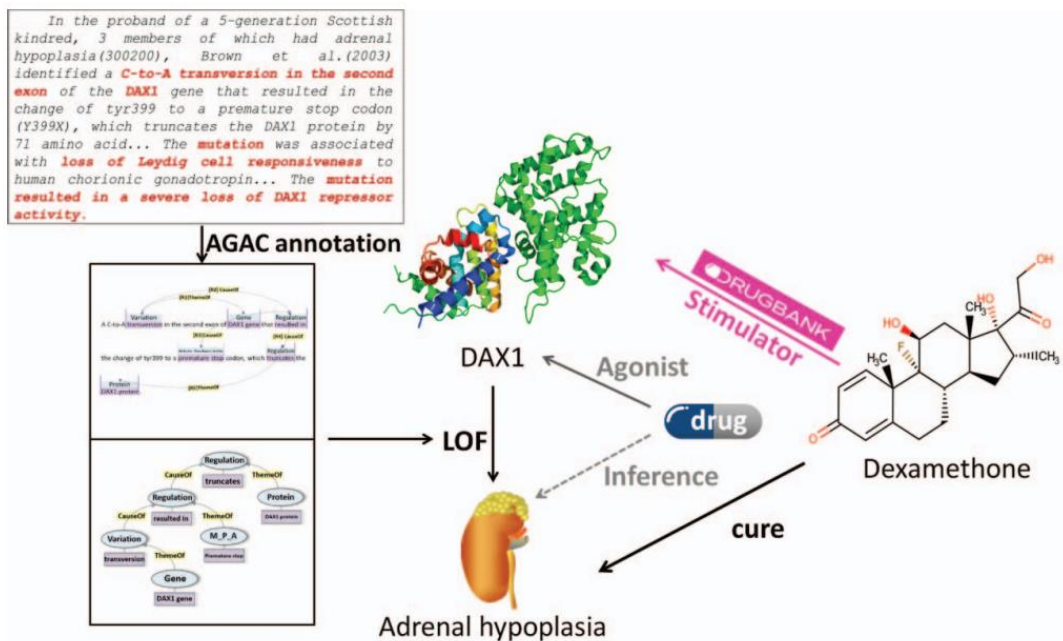


AGAC 语料库设计了 5 个语义触发标签和 3 个根调控触发标签作为分子水平和细胞水平的生物实体标注，重点关注基因突变导致的生物过程中功能变化的信息，上图说明了对生物活动给不同标注的依据。

更具体的标注指南说明见：

Yuxing Wang, Xinzhi Yao, Kaiyin Zhou, Xuan Qin, Jin-Dong Kim, Kevin Bretonnel Cohen, Jingbo Xia*. Guideline Design of an Active Gene Annotation Corpus for the Purpose of Drug Repurposing. 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI 2018), pp:1-5, Oct, 2018, Beijing

3、AGAC 在知识发现的应用 抗癫痫药物发现



上图是“地塞米松”药物发现过程的流程图，通过文本挖掘完全发现了突变 DAX1 的 LOF 功能变化，且没有任何给定药物的先验信息，任何针对 DAX1 的匹配药物的发现都是有价值的。

为了验证这一点，于是在 DrugBank 中搜索 DAX1 的靶向信息，发现了一种名为“地塞米松”的药物，它不仅是一种激动剂，而且靶向 DAX1。

更多 AGAC 在抗癫痫药物中的应用见已发表文献：

Yuxing Wang, Kaiyin Zhou, Jin-Dong Kim, Kevin Cohen, Mina Gachloo, Yuxin Ren, Shanghui Nie, Xuan Qin, Panzhong Lu, Jingbo Xia*. An Active Gene Annotation Corpus and Its Application on Anti-epilepsy Drug Discovery. BIBM 2019: International Conference on Bioinformatics & Biomedicine. Page: 512-519, San Diego, U.S, Nov, 2019.

4、更多 AGAC 相关信息请点击门户页面：<https://hzaubionlp.com/agac/>