A Benchmark of Discovering Drug-Target Interaction from Biomedical Literature

Yutai Hou[1], **Yingce Xia**[2], **Lijun Wu**[2], **Shufang Xie**[2], **Yang Fan**[3],
Jinhua Zhu[3], **Wanxiang Che**[1], **Tao Qin**[2], **Tie-Yan Liu**[2]
[1] Harbin Institute of Technology [2] Microsoft Research
[3] University of Science and Technology of China
[1] {ythou,car}@ir.hit.edu.cn, [3] {fyabc,teslazhu}@mail.ustc.edu.cn
[2] {yingce.xia,lijuwu,shufxi,taoqin,tyliu}@microsoft.com

A Dataset documentation

(1) Our dataset, KD-DTI, is about to speed up the research of discovering drug, target and their interaction from the literature, which is an important topic. Our dataset is built upon DrugBank and TTD. After the confirmation from the owners of DrugBank, to use our dataset, one should register an account for DrugBank to extract the drug and target names related to DrugBank. For TTD, we confirm with the owner, and no license is required.

(2) Any researchers about machine learning, natural language processing, biology and medicine might benefit from our dataset.

(3) Currently, the dataset is only visible to reviewers through the private URL. After the review process, we will release our dataset through Github or a publicly available website. Our dataset will be maintained for a long time.

(4) We have confirmed with the owners of DrugBank and TTD for re-distribution.

(5) The licence of the dataset is the Computational Use of Data Agreement (C-UDA) License.

(6) We will update our dataset regularly according to the feedback of users.

B Detailed patterns for data filtration

The detailed pattern are summarized in Table 1.

C Study on regularization techniques

We explore various combinations of dropout and label smoothing based on Transformer + PubMedBERT-Fuse. The F1 scores of the validation set are reported in Table 2. The best result is achieved when dropout and label smoothing are set as 0.3 and 0.2 respectively. However, the training F1 score is 70.29, which is significantly larger than the validation set. We keep increasing dropout and label smoothing, and found that the validation performance cannot be further improved. This shows that using basic techniques to improve generalization (i.e. dropout, label smoothing) can bring limited improvement, and we need more effective regularization techniques for the DTI triplet extraction task.

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.

Input: A query q and the retrival results $\mathcal{R}(q, D) = \{r_j\}$.	
Reliable patterns	
P1	$\exists r_j \in \mathcal{R}(q, D)$ s.t. r_j and q are exactly the same (except for parentheses, cases, and punctuation marks);
P2	$\exists r_j \in \mathcal{R}(q, D)$ s.t. r_j and q have > 20 characters or ≥ 3 words in common;
Positive patterns	
P3	There are at least 2 elements in $\mathcal{R}$ matching variant formats of q (e.g., +s, +ed, +ing);
P4	There are at least 2 elements $r_\cdot$ in $\mathcal{R}$ s.t each r and q has > 8 characters in common, or at most 10% different characters;
P5	$\lvert\mathcal{R}(q, D)\rvert > 3$;
Negative patterns	
P6	Each r in $\mathcal{R}$ has < 8 characters, or r is a meaningless word like "other", "unknown", etc.

Table 1: Matching patterns between a query and a paper.

(dr,ls)	0.1	0.2	0.3
0.2	18.05	22.24	21.97
0.3	20.06	23.33	21.45
0.4	20.93	20.14	19.86
0.5	18.71	18.97	18.35

Table 2: Results of the F1 on the validation set. `dr` and `ls` stand for dropout and labeling smoothing respectively.

D Main results with standard deviation

Figure 1 presents the standard deviation of results on DrugBank and TTD. On DrugBank, the standard deviation of each model is around 1.5. On TTD, the standard deviation scores are smaller and usually less than 1.0.

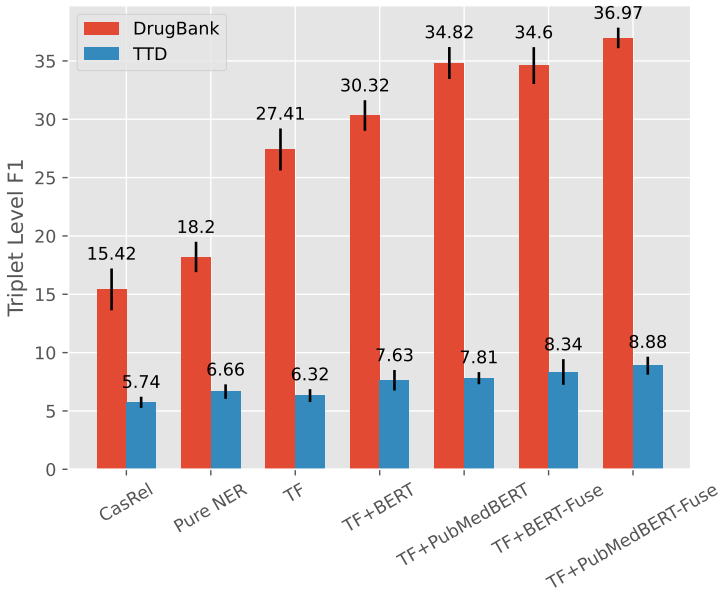

Figure 1: Main results with standard deviation. "TF" denotes Transformer.

Title: Assessment of the abuse liability of ABT-288, a novel histamine H3 receptor antagonist.

Abstract: RATIONALE: Histamine H3 receptor antagonists, such as ABT-288, have been shown to possess cognitive-enhancing and wakefulness-promoting effects. On the surface, this might suggest that H3 antagonists possess psychomotor stimulant-like effects and, as such, may have the potential for abuse. OBJECTIVES: The aim of the present study was to further characterize whether ABT-288 possesses stimulant-like properties and whether its pharmacology gives rise to abuse liability. METHODS: The locomotor-stimulant effects of ABT-288 were measured in mice and rats, and potential development of sensitization was addressed. Drug discrimination was used to assess amphetamine-like stimulus properties, and drug self-administration was used to evaluate reinforcing effects of ABT-288. The potential development of physical dependence was also studied. RESULTS: ABT-288 lacked locomotor-stimulant effects in both rats and mice. Repeated administration of ABT-288 did not result in cross-sensitization to the stimulant effects of d-amphetamine in mice, suggesting that there is little overlap in circuitries upon which the two drugs interact for motor activity. ABT-288 did not produce amphetamine-like discriminative stimulus effects in drug discrimination studies nor was it self-administered by rats trained to self-administer cocaine. There were no signs of physical dependence upon termination of repeated administration of ABT-288 for 30 days. CONCLUSIONS: The sum of these preclinical data, the first of their kind applied to H3 antagonists, indicates that ABT-288 is unlikely to possess a high potential for abuse in the human population and suggests that H3 antagonists, as a class, are similar in this regard.

Prediction: (Drug: "ABT-288", Target: "Histamine H3 receptor", Interaction: "antagonist")

Annotation: (Drug: "ABT-288", Target: "Histamine H3 receptor", Interaction: "antagonist")

Title: Mechanisms of Glucose Lowering of Dipeptidyl Peptidase-4 Inhibitor Sitagliptin When Used Alone or With Metformin in Type 2 Diabetes: A double-tracer study.

Abstract: OBJECTIVE To assess glucose-lowering mechanisms of sitagliptin (S), metformin (M), and the two combined (M+S).RESEARCH DESIGN AND METHODS We randomized 16 patients with type 2 diabetes mellitus (T2DM) to four 6-week treatments with placebo (P), M, S, and M+S. After each period, subjects received a 6-h meal tolerance test (MTT) with [(14)C]glucose to calculate glucose kinetics. Fasting plasma glucose (FPG), fasting plasma insulin, C-peptide (insulin secretory rate [ISR]), fasting plasma glucagon, and bioactive glucagon-like peptide (GLP-1) and gastrointestinal insulinotropic peptide (GIP) was measured.RESULTS FPG decreased from P, 160 $\pm$ 4 to M, 150 $\pm$ 4; S, 154 $\pm$ 4; and M+S, 125 $\pm$ 3 mg/dL. Mean post-MTT PG decreased from P, 207 $\pm$ 5 to M, 191 $\pm$ 4; S, 195 $\pm$ 4; and M+S, 161 $\pm$ 3 mg/dL ($P < 0.01$]. The increase in mean post-MTT plasma insulin and in ISR was similar in P, M, and S and slightly greater in M+S. Fasting plasma glucagon was equal (65-75 pg/mL) with all treatments, but there was a significant drop during the initial 120 min with S 24% and M+S 34% (both $P < 0.05$) vs. P 17% and M 16%. Fasting and mean post-MTT plasma bioactive GLP-1 were higher ($P < 0.01$) after S and M+S vs. M and P. Basal endogenous glucose production (EGP) fell from P 2.0 $\pm$ 0.1 to S 1.8 $\pm$ 0.1 mg/kg min, M 1.8 $\pm$ 0.2 mg/kg min [both $P < 0.05$ vs. P), and M+S 1.5 $\pm$ 0.1 mg/kg min ($P < 0.01$ vs. P). Although the EGP slope of decline was faster in M and M+S vs. S, all had comparable greater post-MTT EGP inhibition vs. P ($P < 0.05$).CONCLUSION SM+S combined produce additive effects to 1) reduce FPG and postmeal PG, 2) augment GLP-1 secretion and β-cell function, 3) decrease plasma glucagon, and 4) inhibit fasting and postmeal EGP compared with M or S monotherapy.

Prediction: (Drug: "Sitagliptin", Target: "Dipeptidyl Peptidase 4", Interaction: "inhibitor")

Annotation: (Drug: "Sitagliptin", Target: "Dipeptidyl Peptidase 4", Interaction: "inhibitor"), (Drug: "C", Target: "C-peptide", Interaction: "part of")

Table 3: Case study of triplet generation on unseen sample.

E Case study

We perform a case study to investigate whether a model trained on the KD-DTI dataset is able to discover unseen Drug-Target-Interaction triplets and handle unseen paper. To achieve this, we train a generative model on KD-DTI and make predictions on unseen samples from another dataset, ChemProt, which contains human annotation of chemical-protein relation (a kind of drug-target interaction).We use Transformer+PubMedBERT-Fuse for case study.

As shown in Table 3, in the first case, the entire triplet is correctly extracted, while both the drug and whole triplet are unseen in KD-DTI. In the second case, we successfully extract one of the two annotated triplets. We attribute the missing of the second triplet to the irregular format of drug "C" and the presence of distracting items, such as "P" and "M".

F Broader impact

We propose a new dataset for biomedical knowledge discovery. We believe that this dataset can speed up the research of bioNLP and machine learning. For negative impact, after the success of automatic knowledge discovery, it might cause some unemployment of the related researchers and engineers.

G Distribution of interactions

In the Figure 2, we present a details statistic of interactions included in the corpus.

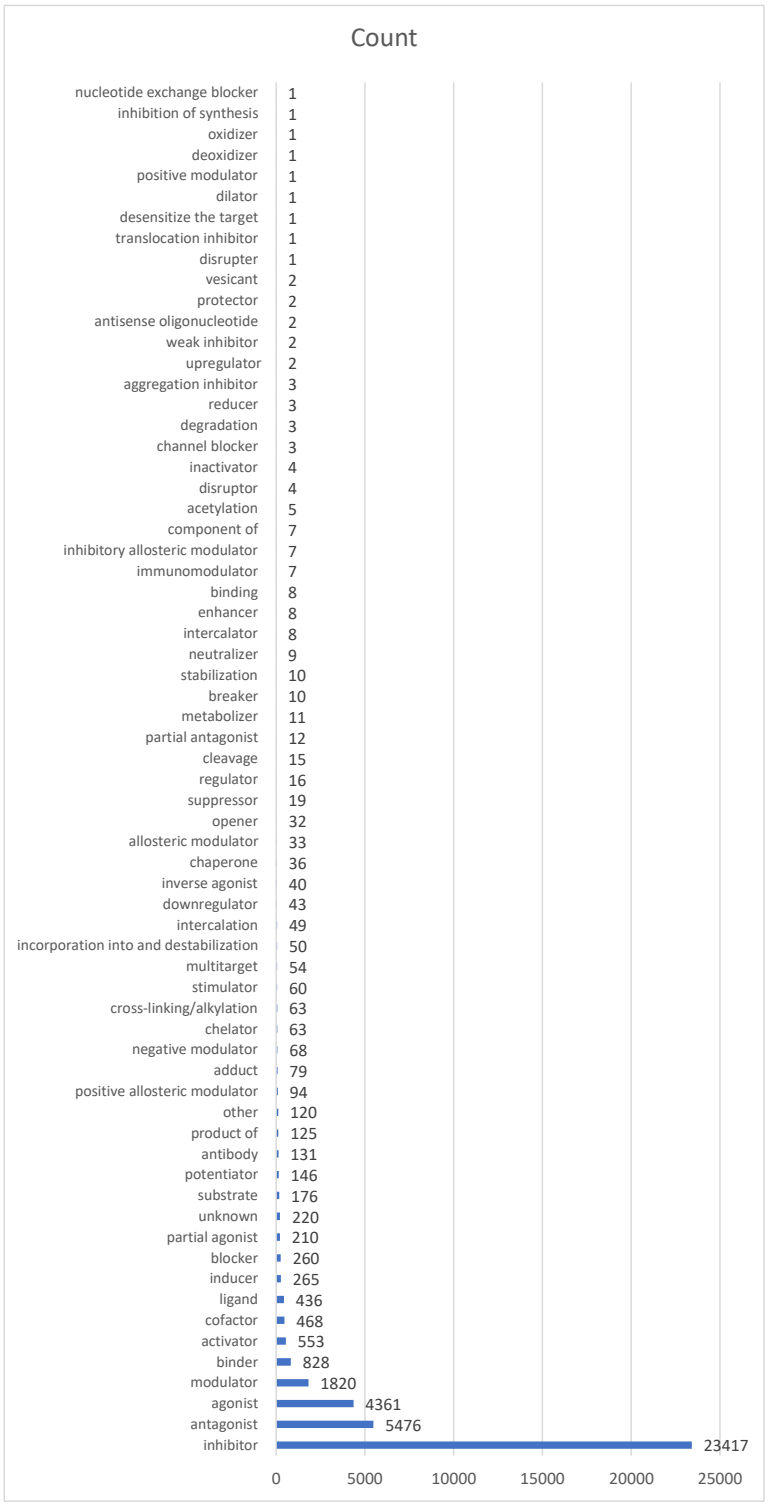

Figure 2: Distribution of interactions.