

Treating COVID-19 with machine learning

Evan Zhang

Dover Sherborn High School, Dover, MA, USA

evanz1005@gmail.com

Abstract. From 2020 to 2023, SARS-CoV-2 destroyed much of our society, while few treatments were available due to the time required for drug discovery. However, with recent advancements in artificial intelligence, it is now ready to fight viruses such as SARS-CoV-2. Chemprop, a machine-learning backbone for molecular properties prediction, can be used to discover novel antiviral drugs by training a classifier model with hundreds of thousands of data points that include molecular information represented by SMILES strings and the observed efficacy in inhibiting SARS-CoV-2 in laboratory tests. The resulting model predicts the effectiveness of untested molecules, which then can be manually tested, minimizing tedious hunting traditionally done by human scientists. With promising performance, the proposed method pushes the boundary of machine learning's involvement in drug research. The trained model achieved a high accuracy in predicting the effectiveness of drugs against SARS-CoV-2 with an AUC score of 0.8455. However, the model loses accuracy when predicting the effectiveness of drugs against SARS-CoV, a different strand of coronavirus, with an AUC of 0.7302. The model was then run on one of the data sets to locate the molecule most likely effective against COVID-19, demonstrating its applicability. The result was a molecule with SMILES string CN1CCN(CC1)C(=O)COC=2C=CC(C)=CC2 also called 1-(4-Methyl-piperazin-1-yl)-2-p-tolyloxy-ethanone. Then the model DrugChat was utilized to determine the properties of the molecule. The model's ability to find likely drugs can hasten drug research drastically, potentially saving countless lives during future pandemics.

Keywords: Machine Learning, Natural Language Processing, Data, Supervised Learning, COVID-19, Drug Discovery, Graph Neural Network

1. Introduction

The COVID-19 pandemic has caused many losses in society, such as its consistent impact on public health [1, 2, 3] and economy [4, 5]. The search for effective treatments [6, 7] against SARS-CoV-2 has been underway since its initial spread. Based on the pathological features and different clinical phases of SARS-CoV-2, many classes of drugs have been developed, such as antiviral agents, inflammation inhibitors, low molecular weight heparin, hyperimmune immunoglobulins, etc. These medicines fight bacterial or viral diseases by either slowing reproductions or killing them[8]. However, the hunt for the perfect antimicrobial drugs that work against a particular disease is difficult and costly due to a pathogen's complex chemical structures and the vast number of candidate drugs to test. Now, years after the outbreak, there are drugs approved for the treatment of COVID-19, such as Nirmatrelvir and Ritonavir, but hastening the drug discovery process would save countless lives [9]. For a much quicker response against emergent diseases, I explored using machine learning algorithms to accelerate drug discovery to counter diseases such as SARS-CoV-2 and tested its potential.

My approach to accelerating drug discovery is to screen promising drug candidates in advance to save time in testing unlikely candidates [10, 11]. Traditionally, this process is performed by human experts or statistical models. With the increasing data sizes and the success of deep learning in computational

biology [12], machine learning, which has strong predictive power, has become a valuable tool for treating newly emerging diseases. Machine learning algorithms use hundreds of thousands of data points of tested substances to capture the essence of an effective drug. Then, after weeding out substances that do not have desirable properties, the remaining substances can be tested much more cheaply.

However, conventional machine learning algorithms like logistic regression and feed-forward networks have been inefficient at drug discovery because they do not adequately utilize the graph structure of molecules. Very recently, a new type of neural network called the graph neural network (GNN) [13] has grown popular since it is much better at processing graphs, leading to several new algorithms for molecular property prediction [14, 15]. The algorithms can take in a molecular structure and predict a substance’s properties.

Still, these algorithms have much more unrealized potential in antimicrobial development. Chemprop¹, a machine learning backbone for antibiotic discovery, has been published in recent years by MIT researchers. It proposes a message-passing neural network to predict molecule properties that can be used for classification. Recently, it has been used to discover Halicin, which inhibits the growth of *E. coli* [15]. However, this tool has not been utilized to its fullest potential against other infectious diseases.

Using Chemprop, data sets regarding the effectiveness of hundreds of thousands of molecules against SARS-CoV-2 could be used to train a classification model to predict the efficacy of a novel substance. The original data is written in Simplified Molecular Input Line Entry System (SMILES), which gives each molecular structure a unique string representation. Thus, data linking particular SMILES strings with their effectiveness in inhibiting SARS-CoV-2 can be used to train a classification model.

In addition, drug effectiveness against COVID-19 is only one of the dimensions that pharmaceutical experts care about. Even if a molecule is effective at inhibiting COVID-19, many other properties, such as the side effects, toxicity, spectrum of treatment, ease of preparation and storage, and mechanism, are also significant in drug discovery with machine learning. The main challenge is considering all these factors in a high-dimensional model. First, predicting all the factors at once is not an easy task because properties such as ease of preparation contain many dimensions and may not have a standard categorization other than through natural language. Secondly, modeling all these factors simultaneously requires a large model capacity, which leads to the leverage of large language models.

The advent of recent large language models (LLM) has revolutionized machine learning through their widespread capabilities. These models, such as OpenAI’s GPT-4, are capable of understanding human-like text with an unprecedented level of sophistication and nuance due to their complexity and training on vast amounts of data. Pre-trained large language models with proper fine-tuning have also demonstrated state-of-the-art performance in domains unrelated to natural languages, such as time series forecasting and image processing. This technology is transforming how we interact with machines. Therefore, I also leveraged the potential of LLMs to discover ideal drugs to treat COVID-19.

In May 2023, DrugChat² [16] emerged as a prototype system using LLMs to answer various questions about drug properties. It is a ChatGPT-like system that uses Vicuna³, an open-source large language model, and a GNN to provide specific information regarding a particular molecule. I used it to analyze the properties of the effective drug candidates my model found, which provides a deeper understanding of how the molecules may behave and why they might be effective.

2. Problem Description

A molecular structure of a molecule can be represented by a graph $G = (V, E)$, which contains atoms as node features $x_v \in V$ and bonds as edge features $e_{vw} \in E$. For a graph, each node and edge can have different values, with there being q nodes and p edges in total. In a data set, this information can be gleaned from the SMILES string. The model passes the string through an algorithm, translating it into

¹ <https://github.com/chemprop/chemprop>

² <https://github.com/UCSD-AI4H/drugchat>

³ <https://lmsys.org/blog/2023-03-30-vicuna/>

a graph, which can then be passed through the Chemprop GNN. The model $F()$ aims to accurately determine whether a particular molecule/graph G will have a specified property. The result is the output $\hat{y} = F(G)$, which identifies whether this molecule is a promising candidate against the virus, where $0 \leq \hat{y} \leq 1$. After finding the most effective molecules, the results can be fed into the DrugChat system for further information about their properties, and a natural language sequence is returned as a reference.

3. Methodology

3.1. Graph Neural Network for Drug Effectiveness Classification

I propose a graph neural network to learn from molecular data and predict the probability of a molecule being effective against the SARS-CoV-2 virus. In a typical graph neural network, each node can obtain and use information from connected nodes to update itself with an aggregation function. In this way, a graph neural network can capture specific patterns in a set of graphs based on their trends. For a molecular graph, edge information is crucial since connections differ between different atoms. Therefore, the updating of both nodes and edges is vital for good performance.

A GNN consists of several layers in which the original graph is mapped into hidden states. The hidden state at step $t + 1$ is updated based on the hidden state at step t until the final output is obtained. At time $t = 0$, the hidden state h_u^0 of node u is the same as the original feature of node x_u . At some step, $t \geq 0$, message passing occurs by creating a message for each node m_u^{t+1} comprising the aggregate of the information from its hidden state h_u^t , all its neighbors' hidden states h_v^t , and the edges between the nodes e_{vu} . All of this information is aggregated with a function $M_t()$, the aggregation function at the t -th layer, where $N(u)$ denotes the set of all nodes connected to node u by edges as neighbors:

$$m_u^{t+1} = \sum_{v \in N(u)} M_t(h_u^t, h_v^t, e_{uv}), \quad (1)$$

This message is then used to modify the node's hidden state from step t to step $t + 1$, h_u^{t+1} :

$$h_u^{t+1} = U(h_u^t, m_u^{t+1}). \quad (2)$$

At the same time, the graph neural network updates edges m_{uw}^{t+1} using a similar method from step t to step $t + 1$:

$$m_{uw}^{t+1} = \sum_{v \in \{N(u) \setminus w\}} M_t(h_{uw}^t, h_{vu}^t, x_u). \quad (3)$$

All edges connected to the edge uw by node u contribute to updating the representation of the edge uw . Once this message is created, the edge's hidden state can similarly be updated based on the message:

$$h_{uw}^{t+1} = U(h_{uw}^t, m_{uw}^{t+1}). \quad (4)$$

There are several choices for the selection of functions $M()$ and $U()$ as long as they can aggregate neighboring information in the graph and update the original node/edge. Here, I chose $M()$ as the mean aggregator:

$$M_t(a, b, c) = \sigma(W_1^t \cdot (a + cW_2^t \cdot b)), \quad (5)$$

W_1 and W_2 are the weights that characterize the importance of neighboring information when updating the representation and are the variables that changed through the model training. σ denotes an activation function that introduces nonlinearity into the model, which is set as the ReLU function $\sigma(x) = \max(0, x)$ in this case.

On the other hand, $U()$ is set as the simple summation of the original feature and the message:

$$U(a, b) = a + b. \quad (6)$$

After repeating the above method T times, a feed-forward neural network $f()$ is adapted to process the updated representation of the molecule graph after T layers:

$$\hat{y} = F(G) = f(h^T), \quad (7)$$

Where $h^T = [h_v^T, \dots, h_N^T, h_{uw}^T, \dots, h_{N'}^T]$ is the concatenation of all the last node and edge hidden states into a vector. Here, a \hat{y} value closer to one represents a higher likelihood of a drug being effective.

Once the model has the classification prediction, it calculates the binary cross-entropy loss function:

$$Loss = -\frac{1}{n} \sum_{i=1}^n y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i), \quad (8)$$

where n denotes the number of data samples in the training data set, y_i denotes the actual label for datapoint i , and \hat{y}_i denotes the predicted label for datapoint i . The optimization objective is to minimize this loss function to update the above model.

The calculation of each layer is shown in Figure 1. The diagram of the molecule on the left serves as an example. There are seven nodes (atoms) and six edges (chemical bonds). The node aggregation for x_0 , as shown in the right of the figure, takes all the nodes with directed edges towards itself as input to update the representation of x_0 . The edge aggregation of e_{06} , similar to the node aggregation, takes all the edges towards x_0 (the start node of e_{06}) as input to update the representation of itself.

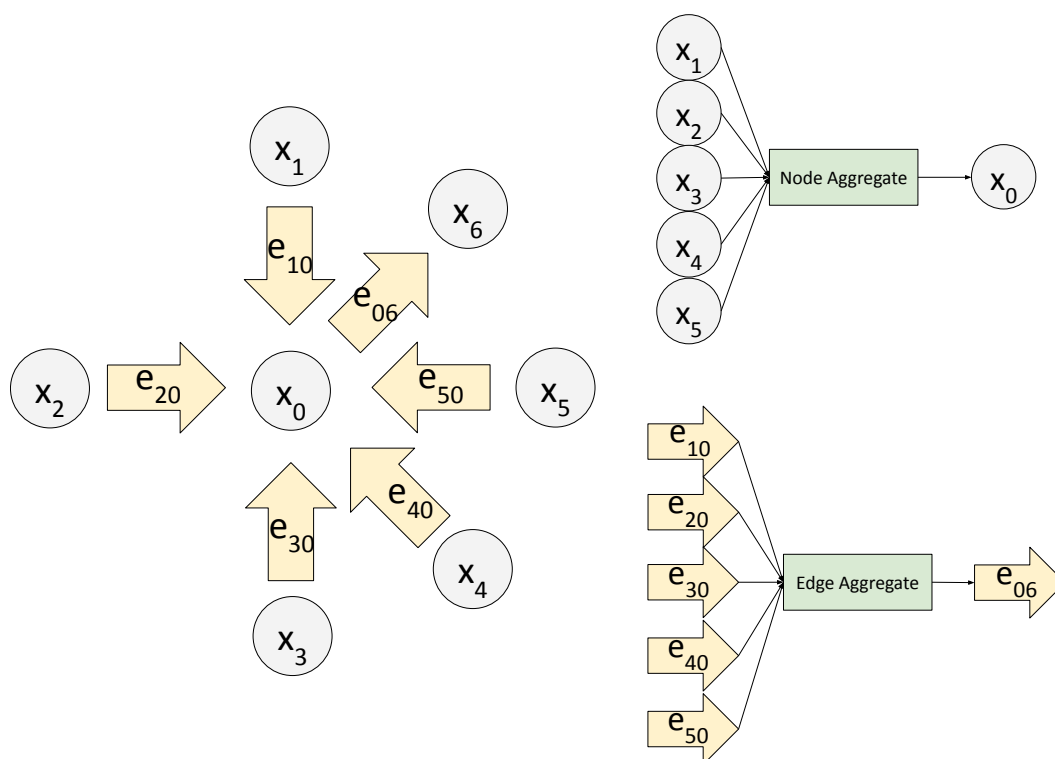


Figure 1. Left: Diagram of Nodes - Right: Node and Edge aggregation.

Moreover, I illustrated the diagram of the full model in Figure 2. The graph neural networks take the molecule graph as input and output the probability of drug effectiveness. After T graph layers, the model flattens the representation of all the nodes and edges into a vector and uses a feed-forward network to process them. After the activation function, it arrives at an estimation of the drug's predicted effectiveness.

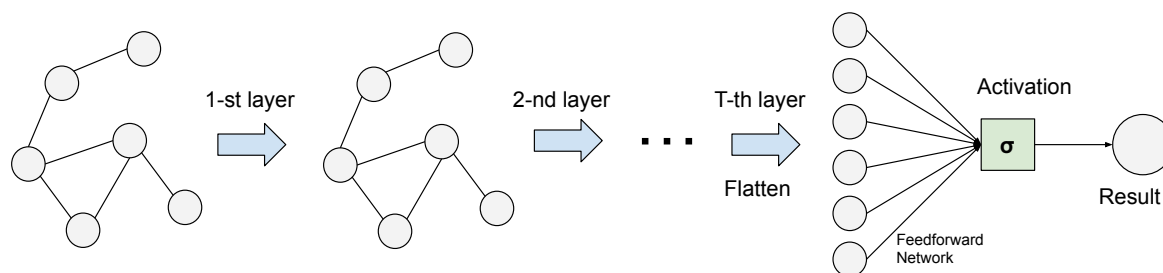


Figure 2. The calculation of graph neural network for drug effectiveness prediction.

The benefit of this model is that it can dynamically update node and edge information simultaneously, compared to a normal aggregation function that uses either node or edge information. This allows for more interactions during learning, increasing the final accuracy and contributing to a more effective drug discovery process.

One critical problem in creating the model is the highly lopsided data set. As one can imagine, not many molecules are effective against COVID-19 when considering random molecule candidates sampled from nature. This data would produce a very lopsided model that would prefer a small number output due to the much higher presence of negative labels in the training data. Therefore, I used a simple yet effective technique to correct the label distribution, i.e., inverse propensity weighting. This means the effective molecules are weighted much more than the ineffective ones to ensure that the model must predict effective molecules correctly to obtain a low loss value.

3.2. Understanding Drug Properties Using Large Language Models

First, I downloaded the Llama model⁴ and Vicuna's delta weights⁵. Then, I constructed the Vicuna model (13 Billion parameters), accompanied by a pre-trained Graph Neural Network and a projection layer to transform the molecular graph into a sequence that the LLM can process.

When DrugChat is fed a SMILES string and a prompt, the SMILES string is first passed through the pre-trained GNN for the model to gather information about the molecule. Then, this information is fed into the Vicuna model with the user-provided question prompt. Then, the Vicuna utilizes a Transformer decoder, which uses a self-attention block that can use tokens to determine the context of different words in the prompt. Vicuna then outputs the information in a string of words that tells the user the answer to the question string. These words are predicted word by word instead of generating whole sentences at once, only utilizing previous words, unlike models such as BERT, which uses all surrounding words. An important note is that the questions asked to Drugchat do not have to be in the training sets. Through this algorithm, a question string and a SMILES string are passed into DrugChat, and an answer is returned. The process is demonstrated in Figure 3.

DrugChat can be used by passing a molecule of interest through the model and asking questions to determine why a molecule has specific properties. Thus, I used the model to explain why some of the molecules my model predicted are effective against COVID-19 and determine some other properties regarding them. While the outputs are not guaranteed to be accurate, it at least gives a glimpse of the mechanisms of a molecule, which is preferable since laboratory testing is very costly and lengthy.

I hoped that with the graph neural network's drug effectiveness prediction and the LLM's understanding of the molecular properties, I could comprehensively prescreen the most promising drug candidates for laboratory testing, resulting in much fewer candidates with a higher probability of being effective.

⁴ https://huggingface.co/docs/transformers/main/model_doc/llama

⁵ <https://huggingface.co/lmsys/vicuna-13b-delta-v0>

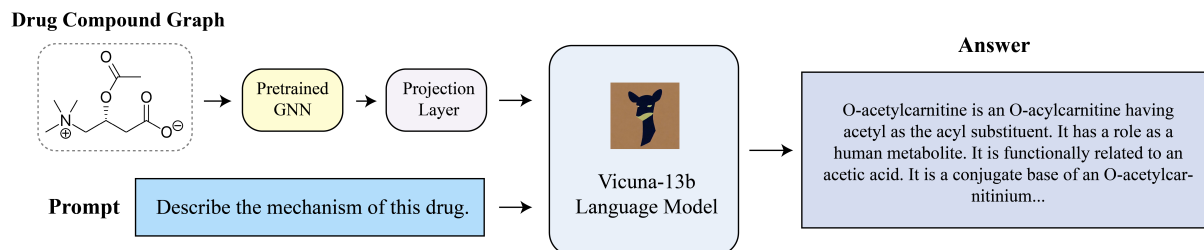


Figure 3. Diagram of DrugChat for understanding the properties of a molecule.

4. Results

4.1. Data Description

My first data set was obtained from the Drug Repurposing Hub⁶, which records efficacy against SARS-CoV-2 via 3C-like protease inhibition, with 405 effective and 290321 ineffective drug results. From this data set, I randomly selected 95% of the molecules for training, with the other 5% being used for testing to determine accuracy. I used the "Main protease structure and XChem fragment screen" data from the Diamond Light Source group⁷ as my second test data set, which measures efficacy like the first data set, with 79 effective molecules and 802 ineffective molecules. I also utilized the PLpro data set (the Bioassay that detects activity against SARS-CoV in yeast models via PL protease inhibition) from PubChem data of AID652038 and AID485353 as the third test data set, which contains 233,891 molecule tests. The training data set and the first test data set are from the same data source, the second test data set is collected from a different data source with different methods of experimentation, and the third test data set was obtained from a similar but different virus, SARS-CoV, which caused the 2002 SARS outbreak. Evaluating different test data sets allows for a better measure of whether the model I developed is generalizable to similar tasks in combating SARS-CoV-2 or its variants. For each datum in the data sets, there is a SMILES string and a boolean label reflecting the efficacy obtained by a human-set threshold.

For the training of DrugChat, two data sets were used: ChEMBL and PubChem. From the ChEMBL data set, 3,892 different drugs were used, paired up with 129,699 question-answer pairs. From the PubChem data set, 6,942 molecules and 13,818 question-answer pairs were used. Some example question-answer pairs are shown below for the molecule with a SMILES string of CC(=O)/N=c1/sc(S(N)(=O)=O)nn1C:

- **Q:** Please provide a description of this drug's mechanism of action.
A: Carbonic anhydrase II inhibitor.
- **Q:** Is this compound an acid, a base, or neutral?
A: Neutral.
- **Q:** How many rotatable bonds does this compound have?
A: 1.
- **Q:** Determine if this drug is administered as a racemic mixture, a single stereoisomer, an achiral molecule, or has an unknown chirality.
A: An achiral molecule.

4.2. Evaluation Metrics

Whether the proposed model effectively screens drug molecules depends mainly on its accuracy, especially in practice. Only an extremely accurate model can speed up drug discovery effectively. Since experimental verification of untested drugs can hardly be done on a large scale and is usually costly, I tested my model

⁶ <https://www.broadinstitute.org/drug-repurposing-hub>

⁷ <https://www.diamond.ac.uk/covid-19/for-scientists/Main-protease-structure-and-XChem>

on existing drug data from different sources to simulate its performance. The accuracy metric can be divided into two dimensions: true positive rate (TPR) and true negative rate (TNR). The true positive rate is the proportion of positive data points correctly identified as positive, while the true negative rate is the proportion of negative data points correctly identified as negative. For this paper, I considered the true positive rate a more important metric to save more time by filtering out undesirable molecules. I also tested my model on the Area-Under-Curve (AUC), which measures a more comprehensive model classification performance with different classification thresholds. Usually, a model with an AUC above 0.8 is considered excellent in classification. The TPR is obtained by selecting the threshold between 0 and 1 that results in the best training accuracy.

4.3. Test Results

Table 1. TPR (True Positive Rate) and AUC (Area Under Curve)

data sets	Metric	The Proposed Model
Training	TPR	93.42%
	AUC	0.9210
First Test	TPR	90.06%
	AUC	0.8957
Second Test	TPR	85.37%
	AUC	0.8455
Third Test	TPR	76.81%
	AUC	0.7302

The main results are shown in Table 1 and Figure 4.

The model performance on the training data set is outstanding, with a TPR of 93.42% and an AUC of 0.921, demonstrating that the proposed model can successfully capture the molecule graph structure patterns. However, the model may contain data sampling bias for better predictions of the training data. Therefore I ran the model on the first test data set to evaluate its performance without the sampling bias, which also did very well, with a TPR of 90.06% and an AUC of 0.8957. This shows that my model performs very well for data from the same source, producing accurate predictions to screen out effective drugs for SARS-CoV-2.

Additionally, it can be noted that before sample reweighing in the initial training of the model, the first test had a 0.8910 AUC. However, after sample reweighing to balance the labels, the AUC increases to 0.8957, demonstrating that the label-balancing strategy effectively improves prediction accuracy.

The second test data set is collected from a different data source. Therefore, the data in the second test data set may vary from the training data set. I tested the model's effectiveness on the data from the second test data set to evaluate its robustness and reliability. Note that it is okay for the model to perform worse on data from different sources since the results of the same drug may vary under biological experimental conditions. This evaluation reveals to what extent the model can be universally effective for the virus in the real world, exempt from the errors and biases that originated from particular experimental processes. I found that the TPR on the second test data set only dropped by about 5%, indicating that the data bias originating from biological experimentation may have affected the machine learning prediction performance.

In light of the rapidly changing variants of the SARS-CoV-2 virus, I was also interested in testing if the machine-learning model could adapt to different viruses or variants. Therefore, I evaluated the model's performance on a third test set, which targeted the SARS-CoV virus, an earlier coronavirus, originating in around 2002. There was a drastic drop in the TPR and AUC, reflecting a decreased prediction accuracy,

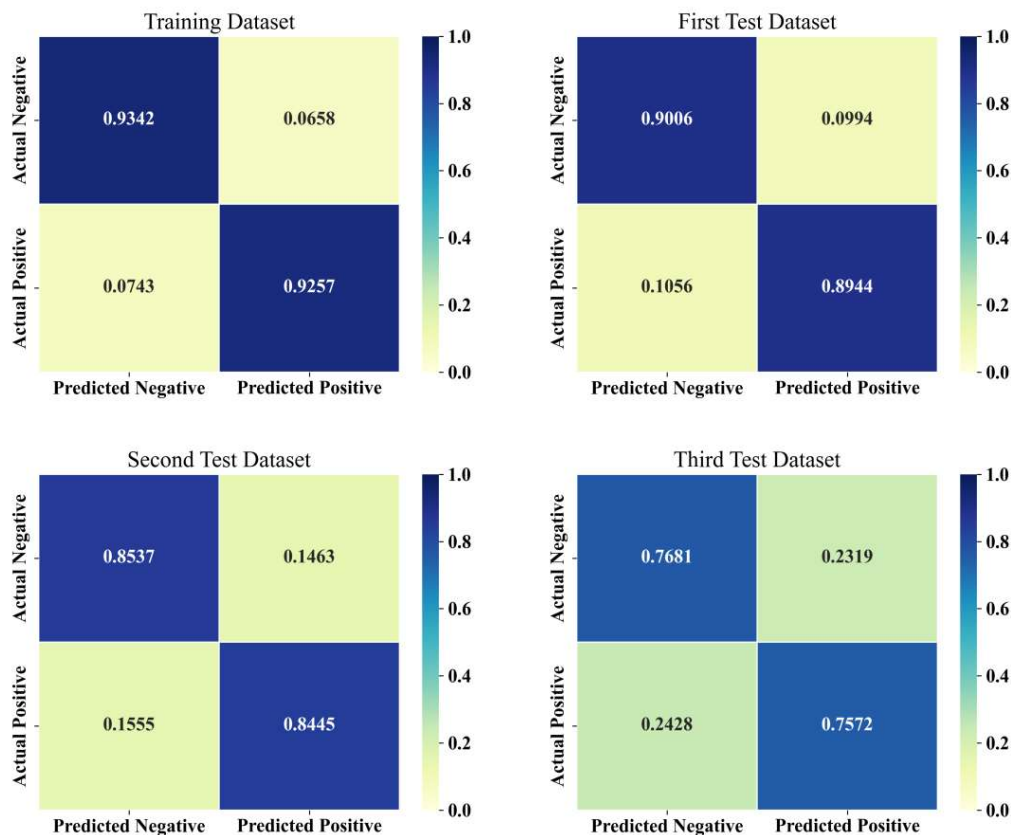


Figure 4. Confusion matrix of the drug effectiveness classification.

with an AUC of only 0.7302. This shows that the model generalization ability is still at risk if the particular virus starts to mutate. When the virus starts to change, the model would also need to change to keep up with the development of the virus. Otherwise, the model will suffer from inaccurate predictions, decreasing its effectiveness in drug discovery.

In addition, I evaluated the time scientists could expect to save for drug discovery. Although it is hard to estimate the time required for high-throughput experiments, my model could obtain predictions in only 0.07 seconds per molecule with the Nvidia A100 graphics card. Since machine learning models can leverage the parallel computation power of the Graphic Processing Units (GPU), they can perform calculations for many drugs simultaneously. Thus, the inference speed for drug discovery is much faster than biological experiments.

4.4. Discovery Example

In the last subsection, I analyze the model's average prediction performance on all data sets. In this subsection, I consider the molecule with the highest effectiveness probability score (\hat{y}) and discuss a potential reason why it was picked. The model can predict whether a candidate molecule is effective against COVID-19 (through the classification) and how confident it is about this discovery (through the output effectiveness probability score). I selected a candidate with the highest effectiveness probabilities from the second test data set, the molecule's SMILES being CN1CCN(CC1)C(=O)COC=2C=CC(C)=CC2. This molecule's scientific name is 1-(4-Methyl-piperazin-1-yl)-2-p-tolyloxy-ethanone. The high output probability of this molecule may be because the model can discern the protease repression mechanism

and how it affects the reproduction of the SARS-CoV-2 virus. The model would need further testing to see its effects on actual SARS-CoV-2 viruses.

Since I could not test my results in a laboratory, I turned to DrugChat to glimpse how the molecule functioned. However, since the technology was only developed in the past year, it is still quite robust, and results are not guaranteed to be accurate. The prompts I used are summarized after "Q:" and the LLM answers are summarized after "A:".

- **Q:** What functional groups does the molecule CN1CCN(CC1)C(=O)COC=2C=CC(C)=CC2 have?
A: The molecule contains a piperazine ring (CN1CCN(CC1)), an ester functional group (C(=O)OC), and a benzene ring (C=CC(C)=CC).
- **Q:** Is molecule CN1CCN(CC1)C(=O)COC=2C=CC(C)=CC2 as a racemic mixture, a single stereoisomer, an achiral molecule, or has an unknown chirality?
A: An achiral molecule. An achiral molecule is one that does not have a non-superimposable mirror image, or in other words, does not exhibit chirality.
- **Q:** What is the molecular formula of this compound, including any salt that it may have?
A: C12H16N2O2.

In the experiments, understanding higher-level properties can be tricky, even for the LLM. To make the model work properly, there would need to be more data for fine-tuning, which I do not have. Some answers from the LLM seem vague and uncertain, which could be both good and bad: while this means that no definite answer is given, meaning that no new information is learned to improve the screening of molecules, it also will not provide misleading answers with certainty. Some further results are shown in the following:

- **Q:** Determine if this compound is a prodrug.
A: From this structure, it's not apparent that the molecule is a prodrug, but it's also not possible to definitively say it is not a prodrug. In the molecule you've mentioned, there is an amide group present. While amide bonds can sometimes be cleaved in the body to activate a prodrug, amide bonds are generally quite stable and not easily cleaved by human metabolic enzymes. Therefore, unless the amide bond in this compound is specifically designed to be metabolized in a particular way, it's unlikely that this compound is a prodrug based on this structure alone. However, determining if a molecule is a prodrug often requires more information than just its structure, such as pharmacokinetic and pharmacodynamic studies, as well as information about the specific metabolic pathways it follows in the body. As such, without more information, it's difficult to definitively determine if this molecule is a prodrug or not.
- **Q:** Is this drug used for therapeutic purposes, rather than for imaging, additives, or other non-therapeutic applications?
A: The molecule described by the SMILES notation "CN1CCN(CC1)C(=O)COC=2C=CC(C)=CC2" is a piperidine derivative with an amide and an aromatic group. Molecules with similar structures are indeed used for therapeutic purposes, such as pharmaceutical drugs. However, whether this particular molecule is used for therapeutic purposes would depend on a variety of factors, including its pharmacological properties, safety profile, and the results of preclinical and clinical trials.
- **Q:** Is this compound an acid, a base, or neutral?
A: The molecule described by the SMILES notation "CN1CCN(CC1)C(=O)COC=2C=CC(C)=CC2" has several functional groups. One of these is an amide group, -C(=O)N-, and the other is an aromatic ring with a methoxy group attached. In addition, there's a tertiary amine group present within the piperidine ring. In terms of acid-base chemistry, the molecule is likely to behave as a base because of the presence of the nitrogen atom in the piperidine ring. Nitrogen atoms with lone pair electrons can act as a base, accepting a proton (H+) from an acid. Amide groups are generally not very basic or acidic due to the resonance stabilization of the nitrogen's lone pair electrons, which are delocalized to the carbonyl oxygen. Thus, the amide functionality in this molecule is unlikely to significantly contribute to its acid-base properties.

5. Conclusion

Finding cures for microbial infections requires scanning through millions of candidates, taking time, money, and resources. While this is happening, people around the world are being infected. I used the Chemprop message-passing neural network backbone to create a model capable of predicting the efficacy of drugs due to its ability to gather much of the information from graph structures.

Chemprop utilizes a message-passing system in which nodes and edges are updated simultaneously based on their closest neighbors. After finishing the message passing, the model extracts the information hidden in the molecule's structure. Then it can be passed through a feed-forward neural network to classify the molecule's efficacy.

The data also has to be formatted correctly before use. The data set contained many more ineffective molecules than effective ones, on the order of hundreds. If the data were used without modification, the model would favor outputting negative results much more than positive ones, resulting in a model that rarely outputs positive results. To solve this, I weighted positive data points much more so that the model would be forced to get them correct for good accuracy. Then a large language model fine-tuned on the Q&A format of a drug property description data set is leveraged to provide further information about promising drug candidates selected by my model.

The testing data shows that the model functions very well when testing data it was trained on and maintained high accuracy even after using different data sets from separate research centers. However, the accuracy fell when testing a data set targeting a different variant of the virus, demonstrating that as viruses evolve and mutate, they can become different enough that the model will no longer be effective. New models must be trained over time to find new effective medicines.

Understanding the effectiveness of machine learning algorithms in the short and long term can allow humanity to use them to accelerate the drug research process, potentially saving countless lives.

References

- [1] Amit Sud, Michael E Jones, John Broggio, Chey Loveday, Bethany Torr, Alice Garrett, David L Nicol, Shaman Jhanji, Stephen A Boyce, Firza Gronthoud, et al. Collateral damage: the impact on outcomes from cancer surgery of the covid-19 pandemic. *Annals of Oncology*, 31(8):1065–1074, 2020.
- [2] Brit Long, William J Brady, Alex Koyfman, and Michael Gottlieb. Cardiovascular complications in covid-19. *The American journal of emergency medicine*, 38(7):1504–1507, 2020.
- [3] Mireille Laforge, Carole Elbim, Corinne Frère, Miryana Hémadi, Charbel Massaad, Philippe Nuss, Jean-Jacques Benoliel, and Chrystel Becker. Tissue damage from neutrophil-induced oxidative stress in covid-19. *Nature Reviews Immunology*, 20(9):515–516, 2020.
- [4] Martin McKee and David Stuckler. If the world fails to protect the economy, covid-19 will damage health not just now but also in the future. *Nature Medicine*, 26(5):640–642, 2020.
- [5] Abel Brodeur, David Gray, Anik Islam, and Suraiya Bhuiyan. A literature review of the economics of covid-19. *Journal of Economic Surveys*, 35(4):1007–1044, 2021.
- [6] Cristina Stasi, Silvia Fallani, Fabio Voller, and Caterina Silvestri. Treatment for covid-19: An overview. *European journal of pharmacology*, 889:173644, 2020.
- [7] Susanna Felsenstein, Jenny A Herbert, Paul S McNamara, and Christian M Hedrich. Covid-19: Immunology and treatment options. *Clinical immunology*, 215:108448, 2020.
- [8] Koichi Yuki, Miho Fujiogi, and Sophia Koutsogiannaki. Covid-19 pathophysiology: A review. *Clinical immunology*, 215:108427, 2020.
- [9] Yvette N Lamb. Nirmatrelvir plus ritonavir: first approval. *Drugs*, pages 1–7, 2022.
- [10] Albert P Li. Screening for human adme/tox drug properties in drug discovery. *Drug discovery today*, 6(7):357–366, 2001.
- [11] György M Keserü and Gergely M Makara. The influence of lead discovery strategies on the properties of drug candidates. *nature reviews Drug Discovery*, 8(3):203–212, 2009.
- [12] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Židek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873):583–589, 2021.
- [13] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? *arXiv preprint arXiv:1810.00826*, 2018.
- [14] Zhongkai Hao, Chengqiang Lu, Zhenya Huang, Hao Wang, Zheyuan Hu, Qi Liu, Enhong Chen, and Cheekong Lee. Asgn: An active semi-supervised graph neural network for molecular property prediction. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 731–752, 2020.

- [15] Jonathan M Stokes, Kevin Yang, Kyle Swanson, Wengong Jin, Andres Cubillos-Ruiz, Nina M Donghia, Craig R MacNair, Shawn French, Lindsey A Carfrae, Zohar Bloom-Ackermann, et al. A deep learning approach to antibiotic discovery. *Cell*, 180(4):688–702, 2020.
- [16] Youwei Liang, Ruiyi Zhang, Li Zhang, and Pengtao Xie. Drugchat: Towards enabling chatgpt-like capabilities on drug molecule graphs. *TechRxiv*, 2023.