BARRIER: Beta-Secretase 1 Reduction for Amyloid Plaque Regulation through Inhibition Exploration and Research

Neel Banga
Dougherty Valley High School
San Ramon, CA, United States
neelbanga21@gmail.com

***Abstract*—Alzheimer's is a brain disorder that disproportionately affects older adults with its primary symptom being severe dementia. Worldwide, over 55 million people have Alzheimer's, with 6.7 million affected individuals living in the USA. Current methods to mitigate the effects of Alzheimer's are insufficient with most drugs (e.g., Memantine, Donepezil, Rivastigmine, etc.) being inconsistent while also causing heavy side effects. In order to address these issues, more drugs need to be tested for viability. To speed up the process, this research proposes AI-based models that can potentially detect which drugs will be able to effectively inhibit the crux of the Alzheimer's pathway, an enzyme named Beta Secretase 1. This study documented the investigation of four AI models—K-Nearest Neighbors (KNN), Random Forest, ChemBERTa, and PubChem10M—and their ability to predict drug efficacy for inhibiting BACE1, a vital target in the Alzheimer's Disease (AD) pathway. These models were trained on the ChEMBL4822 database. The KNN and RandomForest models were traditional descriptor-based models whereas the ChemBERTa and PubChem10M models were fine-tuned transformers. The KNN model showed a strong training performance of (R^2 = 0.6092); this score stayed consistent in the testing phase (R^2 = 0.6210). While having a lower score, the RandomForest model displayed similar consistency in the training (R^2 = 0.5651) and testing phase (R^2 = 0.5605). The ChemBERTa model showed significant improvement from the training phase (R^2 = 0.2641) to the testing one (R^2 = 0.6433), indicating high generalization potential. Similarly, the PubChem10M model exhibited large growth from the training (R^2 = 0.2641) to the testing phase (R^2 = 0.6194). These results highlight the unique strengths of each model and underscore the promising role of AI in AD drug discovery. Future work on the refinement and integration of these models could lead to more effective therapeutic agents for AD.**

Keywords-alzheimer's*; *beta-secretase 1*; *machine learning*; *transformer model*; *drug discovery

I. Introduction

Alzheimer's Disease (AD) is a neurodegenerative brain disorder, a type of brain disorder where cells in the central nervous system either fail to work or exist at all [1]. AD has debilitating symptoms (see Figure 1). In the USA alone, nearly 7 million individuals suffer from ***Alzheimer's***; this number is projected to rise to 13 million by 2050. Worldwide, ***Alzheimer's*** and similar dementia are presumed to affect over 55 million individuals and this number does not seem to be going away anytime soon [2].

Currently, most drugs in the market are unable to inhibit the progression of AD, rather they aim to cope with the effects that come with AD (donepezil [3], rivastigmine [4], memantine [5], etc.). The drugs that are able to inhibit the pathway are often controversial, expensive, and come with heavy side effects like brain swelling and microhemorrhages (memantine, lecanemab [6], etc.). With an unfortunate assortment of drugs that aren't able to completely eradicate the disease nor its effects, it is pivotal to find a drug that can effectively inhibit the spread of AD.

It is known that an overexposure/overproduction of Amyloid Plaques in the brain is synonymous with AD [7]; symptoms such as memory loss, poor judgment, lack of spontaneity, reduced cognitive ability, etc. occur when a plaque buildup is formed. Amyloid plaques are abnormal deposits of amino acid chains known as beta-amyloid peptides (Aβ). These are caused by the incorrect cleavage of the Amyloid Precursor Protein (a type 1 transmembrane protein), powered by ***Beta-Secretase*** 1 (BACE1) [8]—see Figure 2.

It is predicted that ***machine learning*** models trained on molecular descriptors and protein structural features will effectively predict IC50 scores for candidate drugs, aiding in identifying compounds with high efficacy in inhibiting ***Beta-Secretase 1*** (BACE1). This predictive capability directly influences the progression of ***Alzheimer's*** disease by enabling the discovery of potent inhibitors targeting the formation of Beta-Amyloid Peptides.

Amyloid plaques do not paint the complete story, however. Tau, an abundant protein in nerve cells, gives neurodegenerative properties to AD [9]. In a healthy organism (without AD), Tau proteins are primarily responsible for stabilizing microtubules. Tau binds to microtubules, ensuring their stability. It assists in nutrient transport within neurons and plays a role in cell division. Aβ and tau interact early in AD pathogenesis, even before the formation of plaques and tangles. Aβ modulates protein kinases and phosphatases, so an overproduction leads to tau misfolding and hyperphosphorylation. Neurofibrillary tangles form within neurons. These tangles consist of aggregated and hyperphosphorylated tau proteins. The accumulation of neurofibrillary tangles disrupts normal neuronal function. Tau tangles b lock communication between neurons, altering memory, cognition, and other brain functions. Tau-induced damage occurs at the synaptic level, where synapses (connections between neurons) are lost. This contributes to cognitive decline in AD. Acetylcholine, a neurotransmitter that plays a vital role in memory, learning, etc [1], is often unable to reach the brain in the presence of a toxic Tau protein therefore causing an acetylcholine deficiency in the brain and propagates the effects of Alzheimer's. Toxic tau enhances Aβ toxicity via

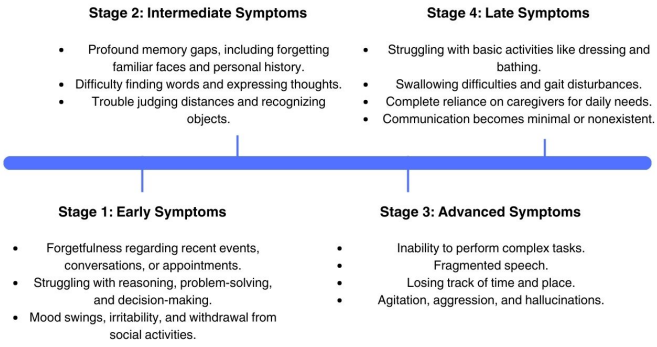

Figure 1. ***Alzheimer's*** Disease Symptoms

a feedback loop, therefore enhancing the symptoms of AD. This leads to a self-propagation of Tau and Aβ (see Figure 2) [10].

Inhibition of BACE1 would cause the suppression of AD due to the absence of Aβ and therefore prevent the formation of a toxic Tau protein. BACE1 is a critical target for the AD pathway as it has an early role in Beta-Amyloid protein production, and produces mild phenotype reactions when deleted, suggesting that inhibiting this enzyme might not have severe side effects and has a well-documented history due to aspartic protease identity. In recent years, there has been a surge in the use of Artificial Intelligence (AI) technology in the medical field. [11]. Various models and architectures have been utilized in biomedical research to enhance its scope and effectiveness. One prominent model that has garnered a lot of attention in recent years is the ***transformer model***. The idea is based on an attention mechanism: a mechanism that allows computers to weigh and understand the context behind different words [12]. This type of model is extremely diverse; it can be used in classification tasks, generative tasks, and even regression tasks [13]. This research does not only focus on the use of transformers; It is important to evaluate multiple models as different models work best for different use cases.

This research is no exception to the use of AI: the goal is to create deep learning models to predict whether drugs can inhibit BACE1 and subsequently find drugs that can disrupt ***Alzheimer's***. For this reason, this project will likely result in an AI model that can accurately identify drugs to inhibit BACE1, as well as find drugs that show large promise to suppress ***Alzheimer's*** [14]. This study explores AI models, including transformers, to predict drugs capable of inhibiting BACE1 and, by extension, tackling Alzheimer's while recognizing that IC50 values are part of a more comprehensive evaluation of drug efficacy. The aim is to develop AI models to identify potential drugs that can effectively inhibit BACE1 and explore promising candidates to mitigate AD progression.

In the related work and methods section, we discuss the related work and methodologies that underpin this study. The methodology section details the datasets and preprocessing techniques employed to prepare our data for analysis. In results,

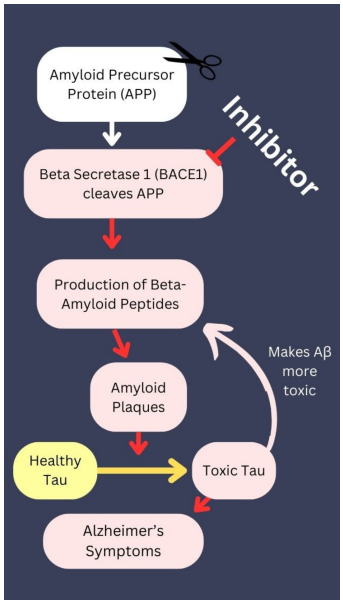

Figure 2. ***Alzheimer's*** Pathway with Tau and BACE1 Inhibition

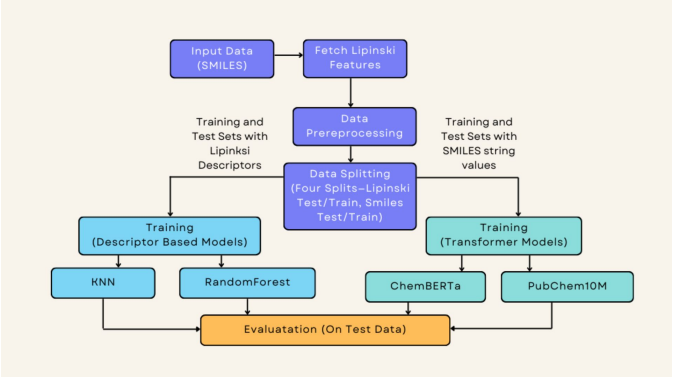

Figure 3. Research Methodology Flowchart

we present the results of our model evaluations, highlighting the performance of each approach. The implications of these results are discussed in the discussion and evaluation portions of this paper, where we also evaluate their significance in the context of Alzheimer's ***drug discovery***. The conclusion concludes the paper, offering insights into future work and potential improvements to the models.

II. Related work | Methods

The fundamental idea this research draws upon is that BACE1 is an effective inhibition target for Alzheimer's reduction. This idea was drawn upon by a previous paper. Gosh et al. proved BACE1 as a potential inhibition target by expressing its numerous advantages: BACE1 is a key target for ***Alzheimer's*** disease (AD) treatment due to its early role in amyloid-β (Aβ) production; the gene deletion of BACE1 produces only mild phenotypes, suggesting that inhibiting this enzyme might not have severe side effects; BACE1 is an aspartic protease, so the mechanism and inhibition of BACE1 are well-documented and

researched, etc [14].

This research also fine-tunes and evaluates transformers that are trained upon molecular properties. Chithrananda et al. built such a model, ChemBERTa [15]. This model takes the well-known RoBERTa transformer [16] and fine-tunes it such that it can predict certain molecular properties. This research takes this model one step further and fine-tunes ChemBERTa to predict whether drugs can effectively inhibit BACE1. The RoBERTa transformer is indeed based on another transformer, BERT [17], which is based on the transformer architecture.

Similar research has been conducted; for example, Baressi et al., conducted research attempting to create models to predict the efficacy of medication on COVID-19 [18]. This research takes that idea further through the evaluation of drugs' efficacy on ***Alzheimer's*** while simultaneously comparing traditional models with newer transformer-based models. This therefore allowed for the evaluation of the difference in efficacy of the two types of models, opening the window for generalizations in this sphere of research.

A. Dataset

To conduct this research, data was collected from the CHEMBL4822 database (Figure 3—input data) that contained different drugs' Simplified Molecular Input Line Entry System (SMILES) notations - simple text-based representations of the drugs molecular structure [19] - paired with their half maximal inhibitory concentration (IC50) scores - a value indicating the dosage needed for a drug to effectively inhibit a certain protein, in this case BACE1. The dataset contains 10619 different drugs (including duplicates) and 46 additional descriptors with the focus being their IC50 scores.

B. Prepossessing

In the preprocessing stage, we first acquired data from the CHEML dataset and filtered it to retain entries where IC50 was specified as the standard type. Null values were addressed with mean imputation, replacing any null values with the dataset's average values, as cited in [20]. We also eliminated duplicate entries, resulting in a refined dataset comprising 7,353 distinct drugs. The data was then narrowed down to the 'canonical_smiles' and 'standard_value' columns. Here, 'canonical_smiles' represents the SMILES notation, and 'standard_value' corresponds to the IC50 value for each drug. To facilitate easier calculations and comparisons, IC50 scores were transformed into pIC50 [18], by taking the negative logarithm of IC50 in molar form. At this stage, our dataset contained two columns: SMILES and pIC50.

Subsequently, we extracted 210 Lipinski features for each SMILES notation entry, thereby expanding the dataset to include 211 columns while maintaining the 7,353 rows. This dataset was then divided into training and test subsets using an 8:2 ratio. For traditional ***machine learning*** models, we utilized the dataset with Lipinski features, whereas for transformer-based models, we retained only the SMILES and pIC50 values. This resulted in four distinct data files comprising training and test sets with Lipinski descriptors as well as training and test sets with SMILES notation. The training sets were employed to develop models using a 5-fold cross-validation approach, while the test sets were reserved to evaluate the models' predictive performance. This careful splitting was important to evaluate how good each model is thoroughly at making predictions.

C. Why Feature Extraction?

This research used two methods to analyze molecular structures: traditional models (KNN and RandomForest) and ***transformer models*** (ChemBERTa and PubChem10M). Traditional models cannot directly process SMILES notation, which represents molecules, ***transformer models*** were also used as they can handle SMILES values directly and may provide better results. For the traditional method, numerical values that describe the drug's properties were needed. 210 descriptors were extracted (such as HeavyAtomMolWt, ExactMolWt, NumValenceElectrons, NumRadicalElectrons, MolWt, etc.), referred to as Lipinski Descriptors [21], from the SMILES strings using the RDKit tool (see Figure 3 to see the data was processed) [22]. This allowed the traditional models to effectively analyze the molecules using these numerical values.

D. Training

After the dataset was both preprocessed and split, AI models were developed to predict drug effectiveness using PIC50 values: two descriptor-based models, KNN [23] and Random Forest [24], and two ***transformer models***, ChemBERTa and PubChem10M [25]).

We selected K-Nearest Neighbors (KNN), Random Forest, ChemBERTa, and PubChem10M models to leverage diverse analytical strengths. KNN and Random Forest are reliable, traditional models ideal for structured data and feature interpretability, providing a solid baseline with molecular descriptors. ChemBERTa and PubChem10M, as transformer-based models, excel in processing sequence data like SMILES strings, capturing complex molecular interactions more holistically. This combination of models allows us to comprehensively evaluate drug efficacy in inhibiting ***Beta-Secretase 1*** (BACE1), balancing robustness with innovative pattern recognition capabilities.

By exploring various hyperparameter settings, each model was trained to find the most optimal configurations. The Mean Squared Error (MSE) [26] and R-squared (R^2) [27] metrics were used to evaluate their performance. The model with the best results, determined by these metrics, was further tested on unseen data to ascertain its R^2 value. This process aimed to find the most suitable model for identifying drugs that might combat Alzheimer's disease, contributing to the discovery of potential treatments and benchmarking different modeling methods. The R^2 metric is a statistical measure often used to assess the accuracy of a regression task. Baressi et al. uses this metric when evaluating the accuracy of AI models to predict pIC50 values of COVID-19 medication [18].

III. Results

When evaluated, the four ***machine learning*** models—K-Nearest Neighbors (KNN), Random Forest, ChemBERTa,

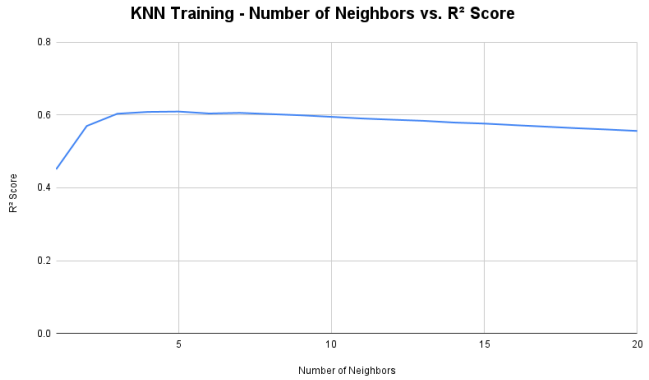

Figure 4. KNN Training Graph - Number of Neighbors vs. R² Score

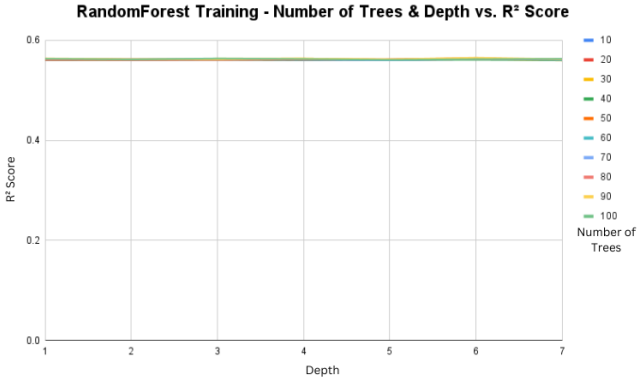

Figure 5. RandomForest Training Graph - Number of Trees and Depth vs. R² Score

and PubChem10M—demonstrated notable differences in their ability to predict certain drugs' inhibition of BACE1. On training data, the KNN model (K=5) was the most effective, boasting the highest R² score of 0.6092 (see Figure 4). It was closely followed by the Random Forest model, with an R² of 0.5651 (maximum depth of 7 and 60 trees) (see Figure 5). In third place was the PubChem10M model (trained for 50 epochs with a learning rate of 0.001), achieving an R² score of 0.4672 (see Figure 6). Finally, in last place for training data, was the ChemBERTa model, with an R² score of 0.2641 (see Figure 7 and Table 1). However, on testing data, the models' results demonstrated a significant shift. The ChemBERTa model led the pack with an R² score of 0.6433, indicating strong generalization to unseen data. This was closely followed by the KNN model with an R² score of 0.621—a continuation of its robust performance in training. The PubChem10M model also showed substantial improvement, achieving an R² score of 0.6194. Lastly, the Random Forest model displayed great consistency, scoring an R² of 0.5605 (see Table 1).

The observed trends underscore the architectural advantages of transformers, such as ChemBERTa and PubChem10M. These models excel due to the attention mechanism, which enables them to capture and generalize complex data patterns inherent

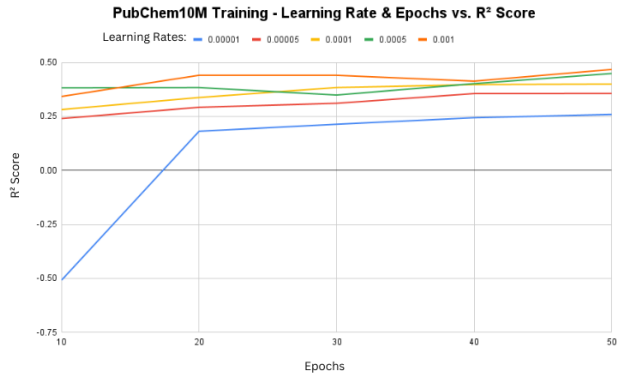

Figure 6. PubChem10M Training Graph - Learning Rate and Epochs vs. R² Score

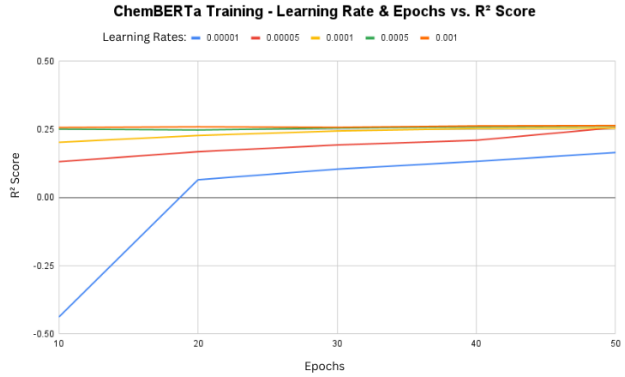

Figure 7. ChemBERTa Training Graph - Learning Rate and Epochs vs. R² Score

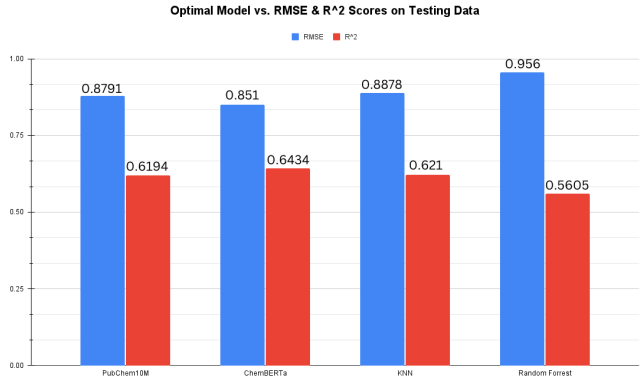

Figure 8. Optimal Model vs. RMSE and R² Scores on Testing Data

TABLE I
RESULT SUMMARY

Model	Training R^2	Testing R^2
KNN	0.609	0.621
Random Forest	0.565	0.560
ChemBERTa	0.264	0.643
PubChem10M	0.467	0.619

in molecular SMILES notation—even when initial training R^2 scores are low. We conducted a series of experiments to better understand these findings, analyzing cross-validation R^2 scores and performance metrics across multiple data subsets. Consistently, the results demonstrated that ***transformer models*** achieve higher test R^2 scores, confirming their superior ability to generalize under varied conditions compared to traditional ***machine learning*** models.

The ChemBERTa model was able to perform well on testing data among metrics boasting the lowest RMSE score of about 0.851 (see Figure 8).

IV. DISCUSSION | EVALUATION

This study contained the evaluation of the KNN, RandomForest, ChemBERTa and PubChem10M models in predicting drugs' ability to inhibit the Alzheimer's pathway using the R^2 score as a metric for the accuracy and effectiveness of each model.

The KNN model performed the best on training data with an R^2 score of 0.6092 with a parameter setting of K=5 neighbors. This indicates the model was proficient in fitting the training data when using K=5 neighbors. In the testing phase, this R^2 score stayed relatively consistent at 0.621.

The Random Forest model had a moderate R^2 score of 0.5651 during training with a maximum depth of 7 and 60 estimators. Like the KNN model, it had a relatively consistent score during testing of 0.5605. This indicates that both the KNN and RandomForest models were relatively consistent models. For the RandomForest model, this consistency might be due to the ensemble nature of Random Forest, which averages multiple decision trees to achieve a robust prediction, making it less likely to be affected by variance in the data [28].

ChemBERTa, using 50 epochs and a learning rate of 0.001, achieved a relatively low R^2 score of 0.2641 during training. That said, the model significantly improved when evaluated on testing data, achieving the highest R^2 score of 0.6433 among all the models. The massive increase in R^2 score suggests that ChemBERTa is particularly good at learning patterns during training and by extension effectively generalizes new data. It shows a large potential for the model to capture the underlying data distribution regardless of initially low training performance.

The PubChem10M model, also trained for 50 epochs with a learning rate of 0.001, had a moderate R^2 score of 0.4672 during training. Like ChemBERTa, the PubChem10M model showed a significant improvement in the testing phase with an R^2 score of 0.6194. This improvement indicates that the PubChem10M model, although not as strong as ChemBERTa, has robust generalization properties. It has comparable testing performance to that of KNN, despite a low initial training score.

Each model showcased unique characteristics across the datasets. The KNN model excelled during training but displayed inconsistencies in testing. The Random Forest model maintained consistent performance across both datasets, but it did not achieve the high levels of accuracy seen in other models. ChemBERTa showed the most notable improvement across the phases. The incorporation of both traditional and newer transformer-based models allows for this research to effectively create generalizations that are lacking in preexisting research.

The ChemBERTa and PubChem10M models, on the other hand, may have built robust patterns during the training phase that were solid and applicable to the testing data. This could be extremely powerful if refined even further.

These results offer valuable insights into the strengths and weaknesses of each model, guiding future research and practical applications where different models may be more suited depending on the context and the nature of the data.

Since ensemble models are hypothesized to have provided consistency and transformers for greater generalization, one could experiment with combining the ensemble nature of the RandomForest model with a transformer to obtain greater results. One could also expand the number of epochs and learning rate values tested to see if there are model configurations that generate better results.

It is important to note that these models do not account for bioavailability, pharmacokinetics, or potential interactions with other drugs, which are critical factors for clinical outcomes. These aspects are essential for understanding how a drug behaves in the body and how effective it will be in real-world scenarios. Consequently, while the models offer insights into drug potential, a more comprehensive approach that includes these factors is necessary for enhancing clinical relevance.

V. CONCLUSION AND FUTURE WORK

This study documented the evaluation of four distinct AI-based models—K-Nearest Neighbors (KNN), Random Forest, ChemBERTa, and PubChem10M—in their ability to successfully estimate how effectively a certain drug could disrupt the Alzhimer's pathway. In this evaluation, results within the testing and training phase were quite varied when evaluating ChemBERTa and PubChem10M models: both models showed low, unfavorable R^2 scores, yet, when these models got to the testing phase, their scores increased by a large margin boasting R^2 scores of 0.6433 (the highest testing score among all the models) and 0.6194 respectively. This indicated unique generalization prowess among the transformer-based models. The descriptor-based models—KNN and RandomForest—on the other hand, were pretty stable; the KNN model had an R^2 score of 0.6092 during training and an R^2 score of 0.621 within testing—indicating powerful consistency. This same trend applied to the RandomForrest model which had a lower R^2 score of 0.5651 but it stayed pretty consistent reaching 0.5605

during testing. These results imply that the transformer-based models had powerful generalization capabilities whereas the descriptor-based models boasted consistency. This highlights the age-old accuracy vs. precision problem which is present in our study today. The main limitation of this study is that it only looks at IC50 values to assess drug potency. While IC50 is important, it does not fully reflect how a drug will work in real life because factors like how the drug is absorbed, distributed, metabolized, and excreted (ADME), and its toxicity also play a role. Our models do not take into account how drugs interact with the body, which might lead to differences between predicted results and actual effects. This study should be seen as a starting point, and future work should include these other factors to make the models more useful for real-world drug development.

REFERENCES

[1] K. Blennow, M. J. de Leon, and H. Zetterberg, "Alzheimer's disease", *The Lancet*, vol. 368, no. 9533, pp. 387–403, 2006.

[2] W. Thies and L. Bleiler, "Alzheimer's disease facts and figures alzheimer's association", *Alzheimers Dement*, vol. 8, no. 2, pp. 131–168, 2012.

[3] J. S. Birks and R. J. Harvey, "Donepezil for dementia due to alzheimer's disease", *Cochrane Database of systematic reviews*, no. 6, 2018.

[4] J. S. Birks and J. G. Evans, "Rivastigmine for alzheimer's disease", *Cochrane Database of systematic reviews*, no. 4, 2015.

[5] B. Jarvis and D. P. Figgitt, "Memantine", *Drugs & aging*, vol. 20, pp. 465–476, 2003.

[6] G. E. Vitek, B. Decourt, and M. N. Sabbagh, "Lecanemab (ban2401): An anti–beta-amyloid monoclonal antibody for the treatment of alzheimer disease", *Expert opinion on investigational drugs*, vol. 32, no. 2, pp. 89–94, 2023.

[7] R. H. Takahashi, T. Nagao, and G. K. Gouras, "Plaque formation and the intraneuronal accumulation of β-amyloid in alzheimer's disease", *Pathology international*, vol. 67, no. 4, pp. 185–193, 2017.

[8] R. J. O'brien and P. C. Wong, "Amyloid precursor protein processing and alzheimer's disease", *Annual review of neuroscience*, vol. 34, no. 1, pp. 185–204, 2011.

[9] E.-M. Mandelkow and E. Mandelkow, "Tau in alzheimer's disease", *Trends in cell biology*, vol. 8, no. 11, pp. 425–427, 1998.

[10] G. S. Bloom, "Amyloid-β and tau: The trigger and bullet in alzheimer disease pathogenesis", *JAMA neurology*, vol. 71, no. 4, pp. 505–508, 2014.

[15] S. Chithrananda, G. Grand, and B. Ramsundar, "Chemberta: Large-scale self-supervised pretraining for molecular property prediction", *arXiv preprint arXiv:2010.09885*, 2020.

[16] Y. Liu *et al.*, "Roberta: A robustly optimized bert pretraining approach", *arXiv preprint arXiv:1907.11692*, 2019.

[11] K. Basu, R. Sinha, A. Ong, and T. Basu, "Artificial intelligence: How is it changing medical sciences and its future?", *Indian journal of dermatology*, vol. 65, no. 5, pp. 365–370, 2020.

[12] A. Vaswani, "Attention is all you need", *arXiv preprint arXiv:1706.03762*, 2017.

[13] J. Born and M. Manica, "Regression transformer enables concurrent sequence regression and generation for molecular language modelling", *Nature Machine Intelligence*, vol. 5, no. 4, pp. 432–444, 2023.

[14] A. K. Ghosh, S. Gemma, and J. Tang, "β-secretase as a therapeutic target for alzheimer's disease", *Neurotherapeutics*, vol. 5, pp. 399–408, 2008.

[17] J. Devlin, "Bert: Pre-training of deep bidirectional transformers for language understanding", *arXiv preprint arXiv:1810.04805*, 2018.

[18] S. Baressi Šegota, I. Lorencin, Z. Kovač, and Z. Car, "On approximating the pic 50 value of covid-19 medicines in silico with artificial neural networks", *Biomedicines*, vol. 11, no. 2, p. 284, 2023.

[19] D. Weininger, "Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules", *Journal of chemical information and computer sciences*, vol. 28, no. 1, pp. 31–36, 1988.

[20] A. R. T. Donders, G. J. Van Der Heijden, T. Stijnen, and K. G. Moons, "A gentle introduction to imputation of missing values", *Journal of clinical epidemiology*, vol. 59, no. 10, pp. 1087–1091, 2006.

[21] H. Mishra, N. Singh, T. Lahiri, and K. Misra, "A comparative study on the molecular descriptors for predicting drug-likeness of small molecules", *Bioinformation*, vol. 3, no. 9, p. 384, 2009.

[22] G. Landrum, "Rdkit documentation", *Release*, vol. 1, no. 1-79, p. 4, 2013.

[23] M. Steinbach and P.-N. Tan, "Knn: K-nearest neighbors", in *The top ten algorithms in data mining*, Chapman and Hall/CRC, 2009, pp. 165–176.

[24] S. J. Rigatti, "Random forest", *Journal of Insurance Medicine*, vol. 47, no. 1, pp. 31–39, 2017.

[25] Y. Omote, K. Matsushita, T. Iwakura, A. Tamura, and T. Ninomiya, "Transformer-based approach for predicting chemical compound structures", in *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, 2020, pp. 154–162.

[26] M. S. Error, "Mean squared error", *MA: Springer US*, pp. 653–653, 2010.

[27] A. Gelman, B. Goodrich, J. Gabry, and A. Vehtari, "R-squared for bayesian regression models", *The American Statistician*, 2019.

[28] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: A review", *Engineering Applications of Artificial Intelligence*, vol. 115, p. 105 151, 2022.