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

Neel Banga

Dougherty Valley High School

Contact email: neelbanga21@gmail.com

Neel Banga

neelbanga21@gmail.com

Neel Banga is a high school student currently attending Dougherty Valley High School, California, USA.

Professional Experience

Method Internship: Design and Engineering consulting firm (method.com)
Summer 2023

Programming Youtube Channel

March 2021 - Present
https://www.youtube.com/@neelbanga—16.3K Subscribers, 95 Videos, 5M+ Views

Science Fair

#1 Math/CS Contra Costa Science Fair, $500 Chevron Award, CSEF Qualification

Coding Projects (https://github.com/neel-banga)

- Creating a Large Language Model (LLM) like Chat GPT
- Generating Music with AI
- Creating an AI-based Chess Engine
- Stock Market Analysis using AI

Hackathons

- Los Altos Hacks: 1st place
- MountainHacks: 2nd place
- CruzHacks: “Best Beginner”

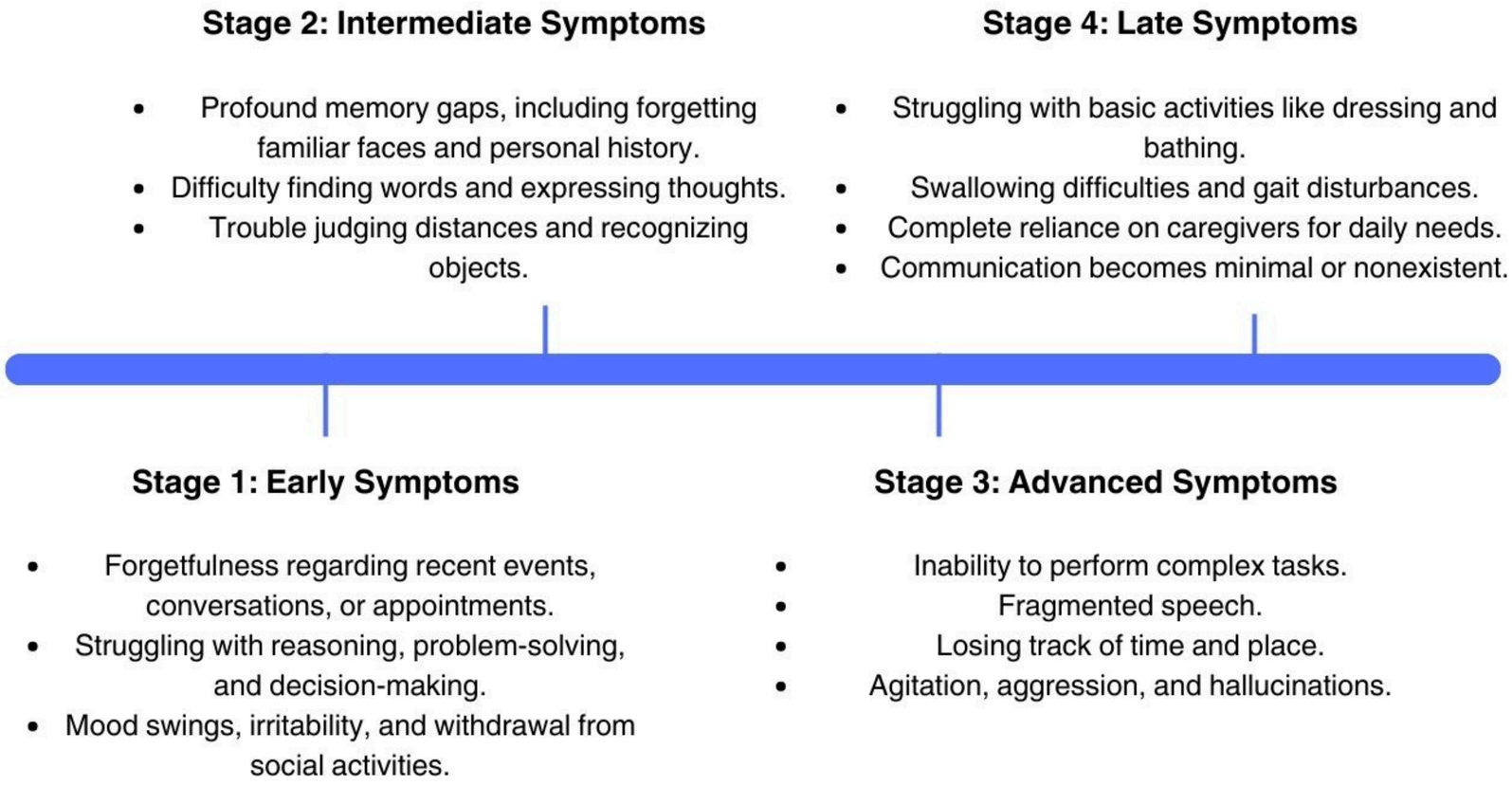
Young Gates

Coding Camp Instructor—taught Python & Scratch to 3rd-5th graders

Alzheimer’s : An Overview

Alzheimer's is a **brain disorder** that **disproportionately affects older adults**; it is imperative to know that Alzheimer's is not a normal part of aging. It has many symptoms with the primary being dementia among affected individuals. **Worldwide, over 55 million people** are affected with this disorder (ADI - dementia statistics), with **6.7 million** of the affected individuals residing **in the United States of America** (2023 Alzheimer's, 2023).

The Alzheimer’s Side Effect Timeline



Current Drugs Used To Treat Alzheimer's

Donepezil

Explanation

Donepezil is an acetylcholinesterase inhibitor (AChE inhibitor). It helps maintain higher levels of acetylcholine in the brain, which is essential for memory and cognitive function.

Side Effects

Nausea, vomiting, diarrhea, loss of appetite, muscle cramps, Insomnia, vivid dreams, and fatigue.

Insufficiency

While donepezil can temporarily improve cognitive symptoms, it does not alter the underlying disease progression. It merely provides symptomatic relief.

Rivastigmine

Explanation

Rivastigmine, like donepezil, is an AChE inhibitor, although minute differences in efficiency, half-life, the method of taking the drug, etc. exist.

Side Effects

Nausea, vomiting, and diarrhea, dizziness, headache, and weight loss.

Insufficiency

Similar to donepezil, rivastigmine only manages symptoms without halting disease progression.

Memantine

Explanation

Memantine is an NMDA receptor antagonist. It regulates glutamate activity in the brain.

Side Effects

Dizziness, headache, constipation, hallucinations and confusion

Insufficiency

Memantine may slow cognitive decline slightly, but it doesn't cure or reverse Alzheimer's.

Memantine

Explanation

Aducanumab targets amyloid plaques, a key feature in the Alzheimer's pathway.

Side Effects

Brain swelling (edema) headache, microhemorrhages (tiny brain bleeds).

Insufficiency

Aducanumab's approval has been controversial due to mixed study results. It's expensive, and its long-term efficacy remains uncertain.

Most of the drugs currently on the market are unable to inhibit the progression of Alzheimer's, rather they aim to cope with the effects that come with Alzheimer's (donepezil, rivastigmine, memantine, 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, etc.).

“The Trigger”: Amyloid Plaques

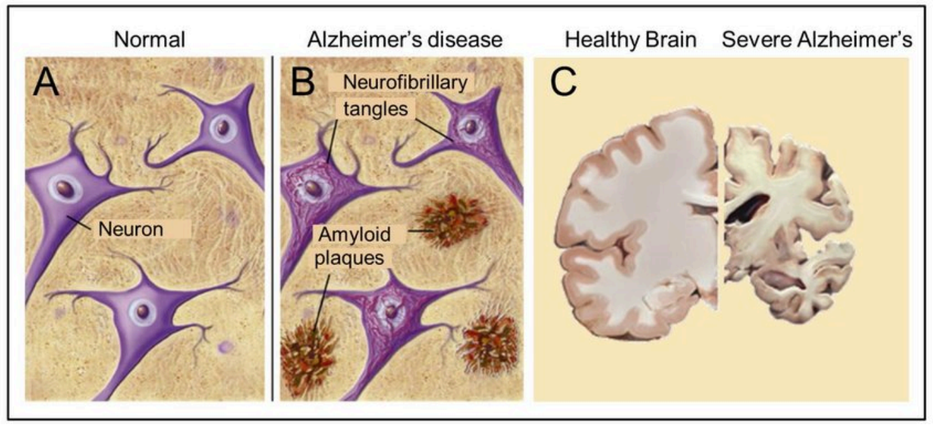
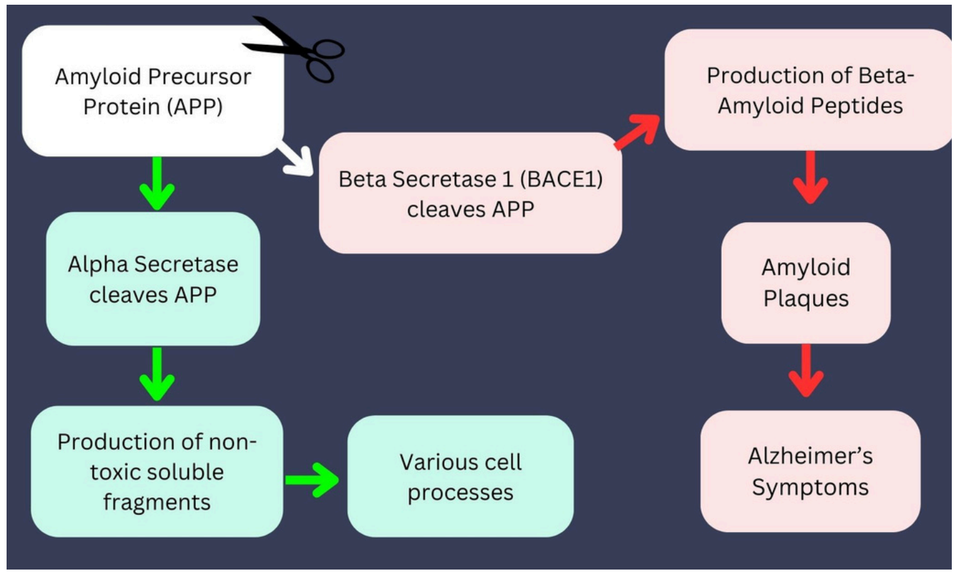
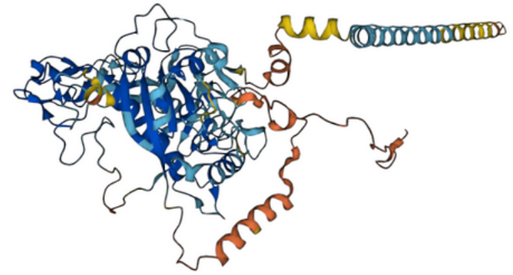


Image taken by 2019 paper Alzheimer’s Disease: Is a Dysfunctional Mevalonate Biosynthetic Pathway the Master-Inducer of Deleterious Changes in Cell Physiology? (Loof, et al.)

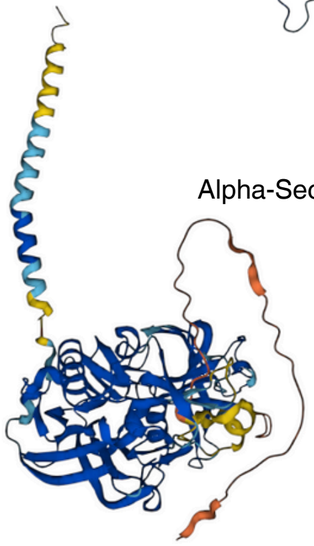


Alzheimer’s pathway (simplified) with an emphasis on Amyloid Plaques

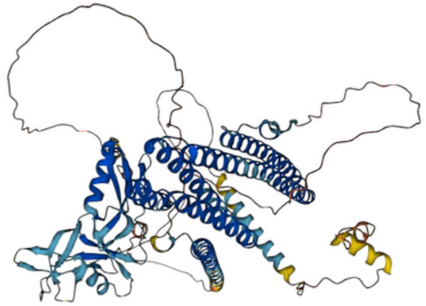
An **overexposure/overproduction of Amyloid Plaques in the brain is synonymous with Alzheimer’s disease (AD)** (Yang Y et al.). Research from the University of Pennsylvania writes that symptoms such as memory loss, poor judgment, lack of spontaneity, reduced cognitive ability, etc. occur **when a plaque buildup is formed.**



Alpha-Secretase 3D Visual powered by Google AlphaFold



Beta-Secretase 1 3D Visual powered by Google AlphaFold

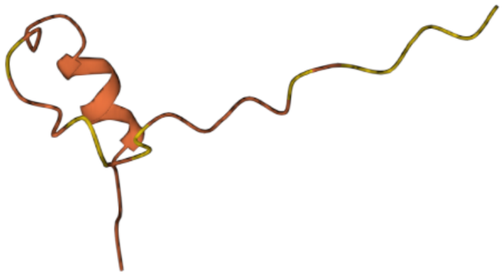


Beta-amyloid protein (aggregation of Aβ) 3D Visual powered by Google AlphaFold

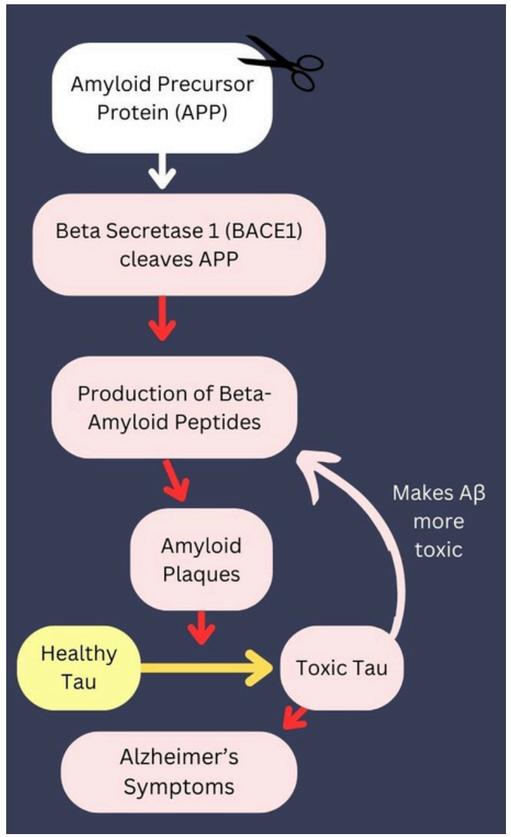
The Amyloid Plaque Pathway

Amyloid plaques **are** simply **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).

“The Bullet”: Tau



Tau protein 3D Visual powered by Google AlphaFold



Alzheimer’s Pathway: Tau & Aβ work in tandem with one another

Tau’s in a Healthy Brain

- In a healthy organism (without AD), Tau proteins are abundant in nerve cells (neurons) and 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.
- In the human brain, tau proteins exist in six isoforms, varying in length from 352 to 441 amino acids. These isoforms differ due to the presence of zero, one, or two inserts of 29 amino acids at the N-terminal portion (exons 2 and 3) and three or four repeat-regions at the C-terminal part (exon 10)

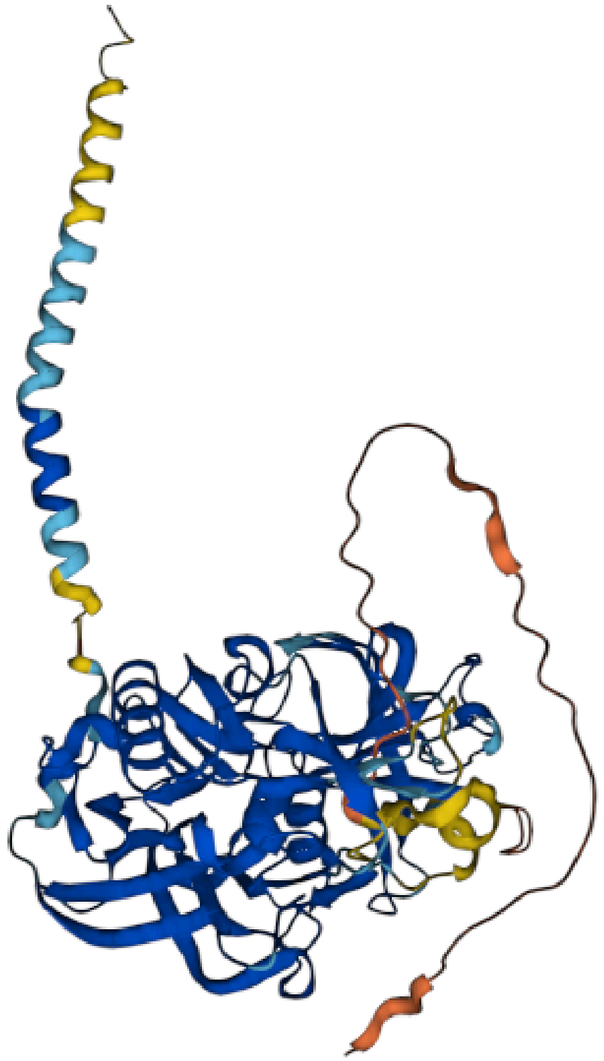
Tau Becomes Toxic

- 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 within neurons form.** These tangles consist of aggregated and hyperphosphorylated tau proteins.
- The accumulation of **neurofibrillary tangles disrupts normal neuronal function.** Tau tangles **block communication between neurons**, impairing memory, cognition, and other brain functions.
    - Eg. Acetylcholine, a neurotransmitter that plays vital a role in memory, learning, etc. is often unable to reach the brain in the presence of a toxic Tau protein therefore causing an **acetylcholine deficiency** in the brain and **propagating the effects of Alzheimer’s.**
- Tau-induced damage occurs at the synaptic level, where synapses (connections between neurons) are lost. This contributes to cognitive decline in AD.
- Abnormal tau proteins become toxic and contribute to neuronal dysfunction. They can trigger acute neuron death and synaptic dysfunction independently of amyloid plaques.

Tau & Aβ Feedback Loop

- **Toxic tau enhances Aβ toxicity via a feedback loop** therefore enhancing the symptoms of AD.
- This leads to a self-propagation of Tau and Aβ.

“The Target” - Beta-Secretase 1

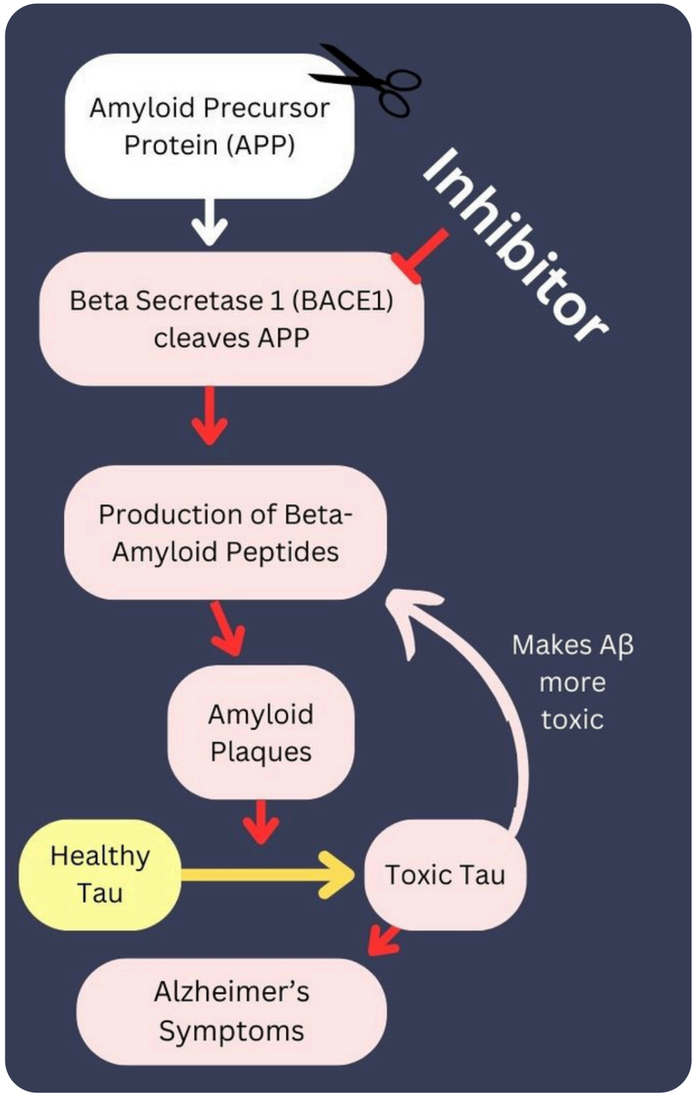


Beta-Secretase 1 3D Visual powered by Google AlphaFold

A 2008 paper published in the Neurotherapeutics journal, written by Arun K. Ghosh, Sandra Gemma, and Jordan Tango research claims that Beta-Secretase 1 (BACE1) is an attractive target for Alzheimer’s pathway inhibition because:

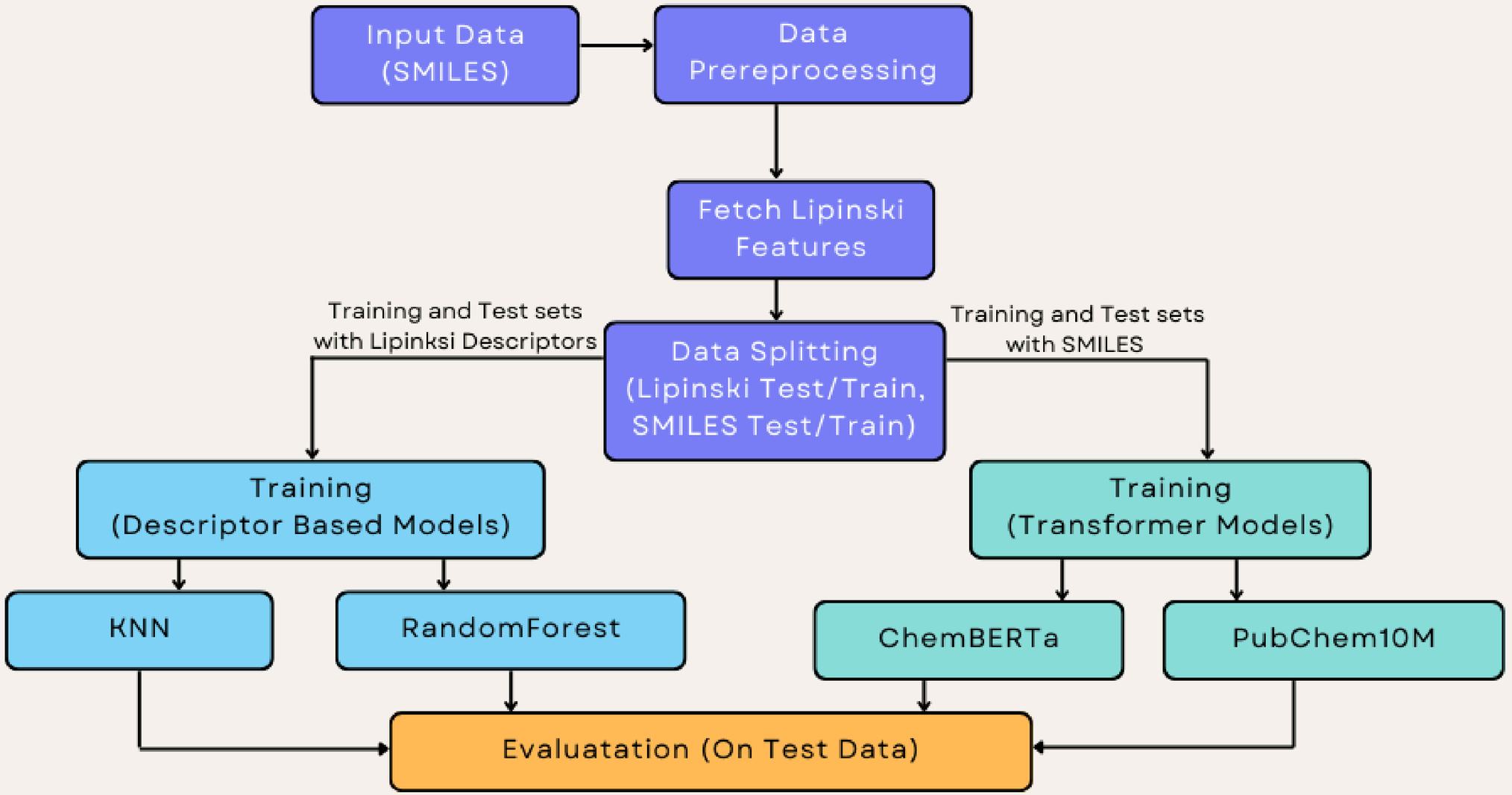
- BACE1 is a key target for Alzheimer's disease (AD) treatment due to its *early* role in amyloid-β (Aβ) production.
- The gene deletion of BACE1 produces only mild phenotypes, suggesting that inhibiting this enzyme might not have severe side effects.
- Due to the fact that BACE1 is an aspartic protease, the mechanism and inhibition of BACE1 are well-documented and researched.

These all point to the fact that BACE1 is a viable target for drugs to approach. Therefore, BACE1 will be the candidate molecule for the research past this slide to be based upon.



Alzheimer’s Pathway: Tau & Aβ work in tandem with one another ***with an emphasis on the Beta-Secretase 1 inhibition***

Approach



Steps taken in building neural networks to effectively speed up drug discovery to inhibit BACE1

Machine Learning Models Used

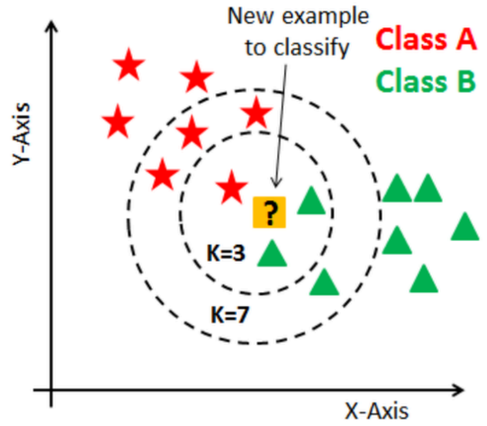


Photo by Sarang Anil Gotke on Kdnuggets

The KNN model plots all the data points on a graph, then simply maps how many "neighbors" the x value in question has, and then takes the average of all the neighbors. **The number of neighbors is a tunable hyperparameter.**

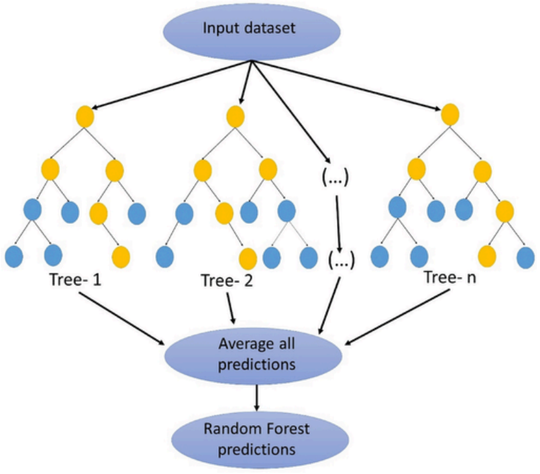


Photo by Sahour et al.

Random Forest is an ensemble learning algorithm that builds multiple decision trees during training and combines their predictions to improve accuracy. It randomly selects subsets of data and features for each tree and then aggregates their predictions to make a final decision. **These decision trees are not constant but rather updated in training to find the most optimal decision tree.**

Although this is a regression task, after doing light testing with logistic/linear regression, they performed too badly during early testing to move to heavy training.

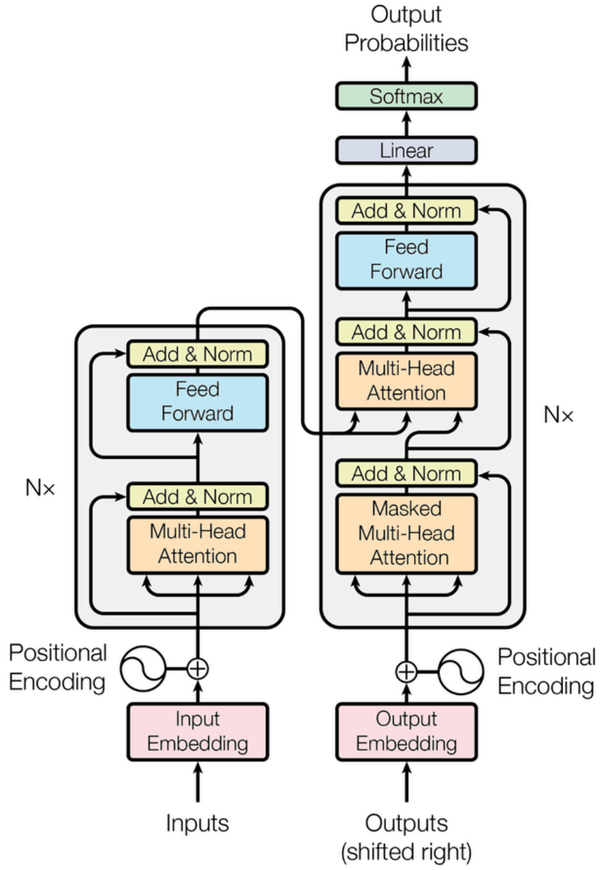


Photo by Dr. Ernesto Lee

The ChemBERTa and PubChem10M models are both trained transformers. Transformers are deep, complex neural networks that can understand the relationships between different encodings. **Just as Chat GPT can understand your English** request, **these models can understand molecular SMILES notation**, the context behind them, and how each element interacts with one another. Using this, the model is able to generate a result, the predicted pIC50 value.

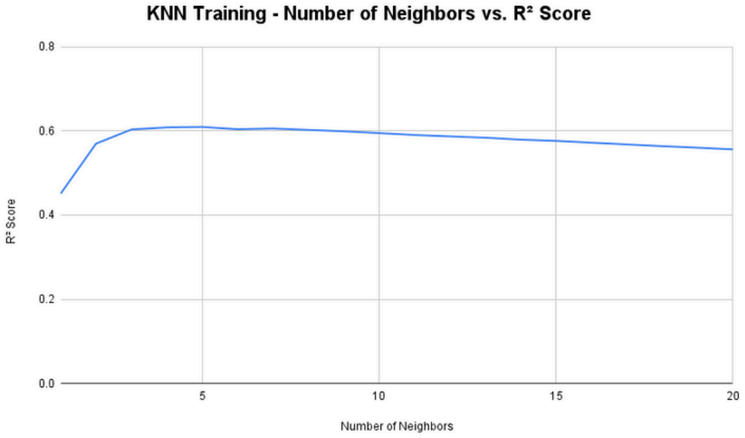
Training The Models

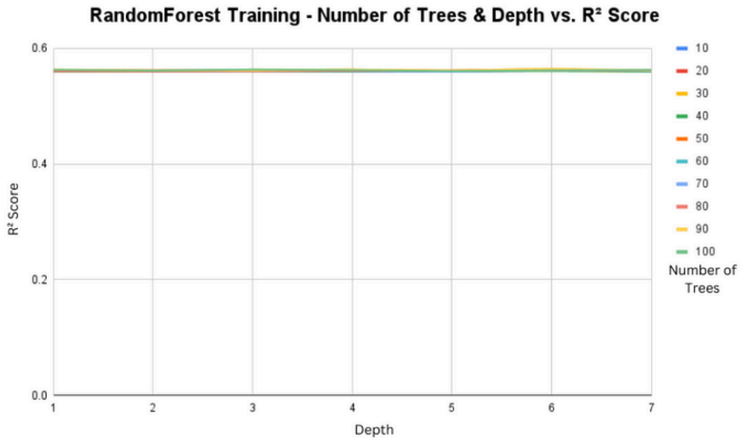
The different algorithms were **trained on the pIC50 scores** and corresponding **SMILES / Lipinski values** (covered in the “Approach” section). Algorithms were trained and observed upon many different hyperparameters to determine which best fit each model. The graphs of this training process are shown below.

The **R^2 score** is **how fit the model is on the training data.** So, we should strive for a **high R^2 score.**

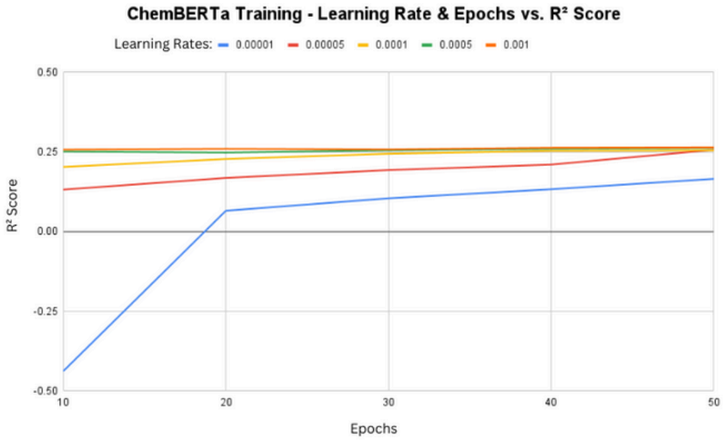
The best-fit models were sent to be tested and compared on new data.

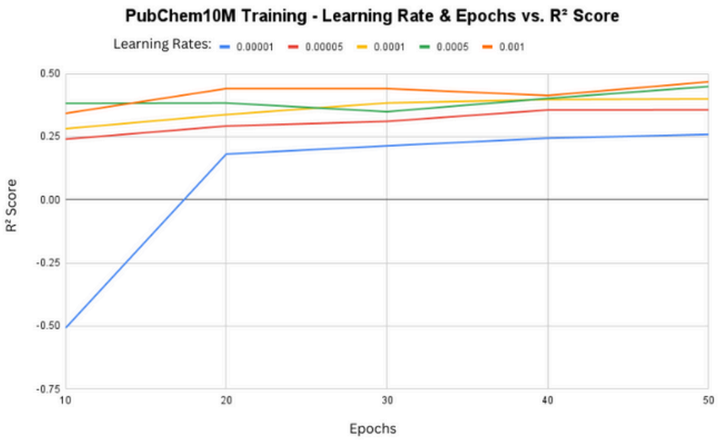
Descriptor-Based Models





Transformer-Based Models

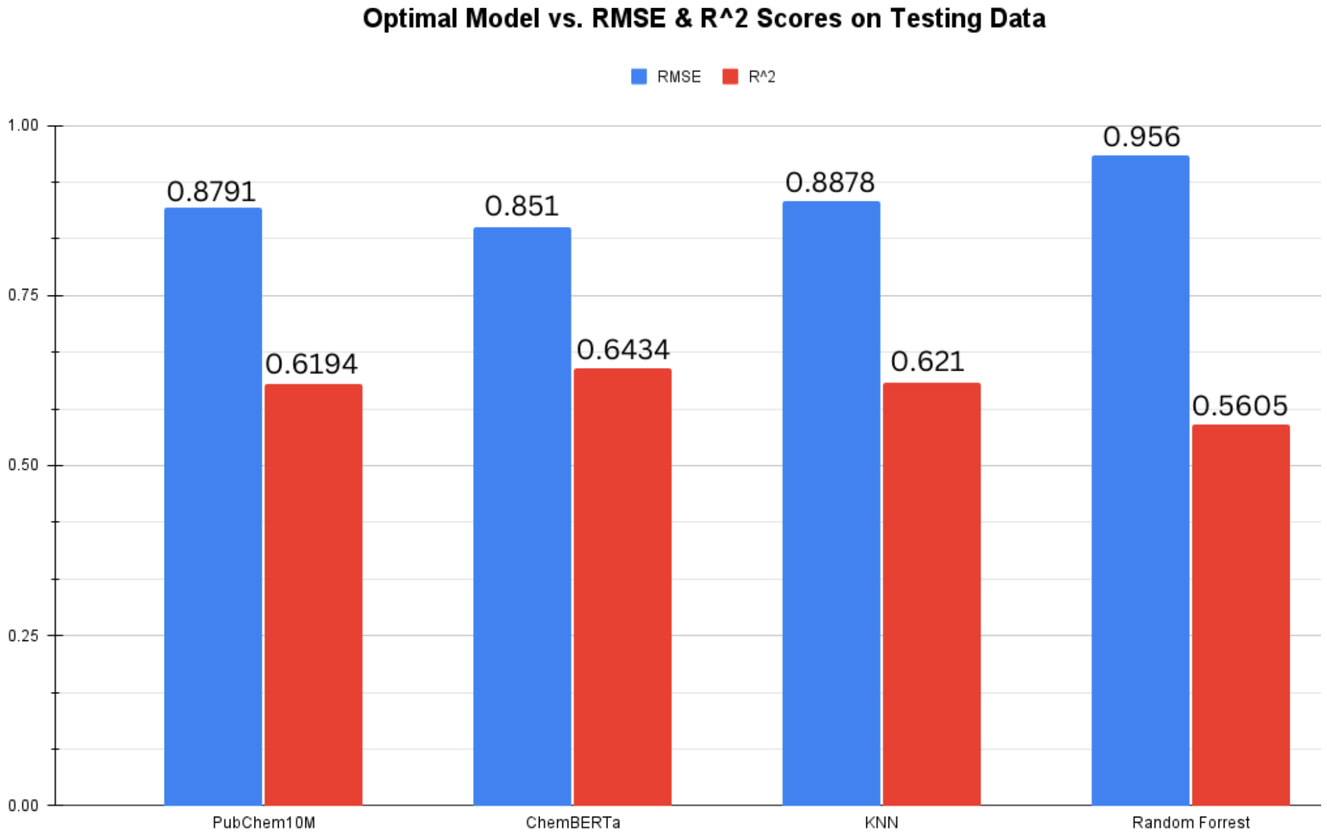




Testing The Models / Discussion

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

While the descriptor-based models maintained consistency through both the training and testing phases. The transformer models, on the other hand, had quite a different story. A drastic improvement was observed for both the ChemBERTa and PubChem10M models in the testing phase—indicating strong generalization capabilities for the two transformers.



The fine-tuned ChemBERTa model (trained with 50 epochs and a learning rate of 0.001) outperformed the other models. With its high R^2 score of 0.643, ChemBERTa shows statistical significance. The PubChem10M model followed closely. The ChemBERTa and PubChem10M models 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.

Conclusion / Future Work

Study Overview

- 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
- The ChemBERTa and PubChem10M models showed low, unfavorable R^2 scores, yet, when these models got to the testing phase, their scores increased by a large margin
    - This pointed to the unique generalization abilities of transformer-based models
- The descriptor-based models—KNN and RandomForest—on the other hand, were pretty stable

Limitations

- Focused solely on IC50 values for drug potency assessment
    - Did not account for ADME (absorption, distribution, metabolism, excretion) or toxicity, which are critical for real-world drug efficacy

Future Work

- While the models offer insights into drug potential, a more comprehensive approach that includes these factors is necessary for enhancing clinical relevance—future research should aim to address this.

Works Cited

[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 ββ-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.

Works Cited

[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-ββ and tau: The trigger and bullet in alzheimer disease pathogenesis", JAMA neurology, vol. 71, no. 4, pp. 505-508, 2014.
[11] S. Chithrananda, G. Grand, and B. Ramsundar, "Chemberta: Large-scale self-supervised pretraining for molecular property prediction", arXiv preprint arXiv:2010.09885, 2020.
[12] Y. Liu et al., "Roberta: A robustly optimized bert pretraining approach", arXiv preprint arXiv:1907.11692, 2019.
[13] 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.
[14] A. Vaswani, "Attention is all you need", arXiv preprint arXiv:1706.03762, 2017.
[15] 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.
[16] A. K. Ghosh, S. Gemma, and J. Tang, "ββ-secretase as a therapeutic target for alzheimer's disease", Neurotherapeutics, vol. 5, pp. 399-408, 2008.

Works Cited

[17] J. Devlin, "Bert: Pre-training of deep bidirectional transformers for language understanding", arXiv preprint arXiv:1810.04805, 2018.
[18] S. Baressi Segota, I. Lorencin, Z. Kovac, 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.

Works Cited

[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.