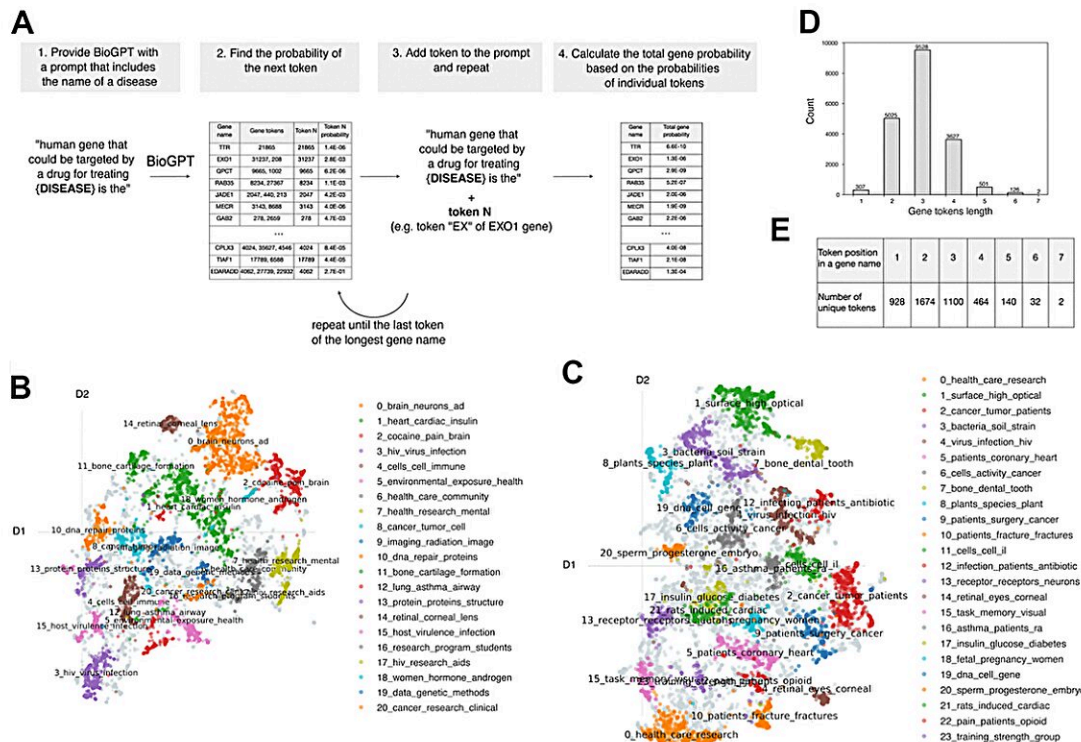# Team trains AI model for age-related disease target discovery

October 4 2023

The main method utilized in the work. Credit: *Aging* (2023). DOI: 10.18632/aging.205055

A new research paper titled "Biomedical generative pre-trained based transformer language model for age-related disease target discovery" has been published in *Aging*.

Target discovery is crucial for the development of innovative therapeutics and diagnostics. However, current approaches often face limitations in efficiency, specificity, and scalability, necessitating the exploration of novel strategies for identifying and validating disease-relevant targets. Advances in natural language processing have provided new avenues for predicting potential therapeutic targets for various diseases.

In their new study, researchers Diana Zagirova, Stefan Pushkov, Geoffrey Ho Duen Leung, Bonnie Hei Man Liu, Anatoly Urban, Denis Sidorenko, Aleksandr Kalashnikov, Ekaterina Kozlova, Vladimir Naumov, Frank W. Pun, Ivan V. Ozerov, Alex Aliper, and Alex Zhavoronkov from Insilico Medicine present a novel approach for predicting therapeutic targets using a large language model (LLM).

"We trained a domain-specific BioGPT model on a large corpus of biomedical literature consisting of grant text and developed a pipeline for generating target prediction," the researchers explain.

This study demonstrates that pre-training of the LLM model with task-specific texts improves its performance. Applying the developed pipeline, the researchers retrieved prospective aging and age-related disease targets and showed that these proteins are in correspondence with the database data. Moreover, they propose CCR5 and PTH as potential novel dual-purpose anti-aging and disease targets which were not previously identified as age-related but were highly ranked in their approach.

"Overall, our work highlights the high potential of transformer models in novel target prediction and provides a roadmap for future integration of AI approaches for addressing the intricate challenges presented in the biomedical field," the team summarizes.

More information: Diana Zagirova et al, Biomedical generative pre-trained based transformer language model for age-related disease target discovery, *Aging* (2023). DOI: 10.18632/aging.205055