Contribution ID: **19**   Type: **Lightning Talk**

Genomic Interpreter: A Hierarchical Genomic Deep Neural Network with 1D Shifted Window Transformer

Tuesday 26 September 2023 17:35 (5 minutes)

Given the increasing volume and quality of genomics data, extracting new insights requires efficient and interpretable machine-learning models. This work presents Genomic Interpreter: a novel architecture for genomic assay prediction. This model out-performs the state-of-the-art models for genomic assay prediction tasks. Our model can identify hierarchical dependencies in genomic sites. This is achieved through the integration of 1D-Swin, a novel Transformer-based block designed by us for modelling long-range hierarchical data. Evaluated on a dataset containing 38,171 DNA segments of 17K base pairs, Genomic Interpreter demonstrates superior performance in chromatin accessibility and gene expression prediction and unmasks the underlying 'syntax' of gene regulation. On the efficiency side, 1D-Swin has time complexity of $O(nd)$, where n is the size of input sequences, d, the window size, is a hyperparameter. This makes it feasible to deal with long-range sequences in other domains, such as Natural Language Processing (NLP) and Time Series Data.

While this work has been presented in the ICML 2023 workshop on Computional Biology, we are actively pursuing collaborations to further advance its practical applications. We make our source code for 1D-Swin publicly available at https://github.com/Zehui127/1d-swin.

Primary authors: LI, Zehui; Dr DAS, Akashaditya; Mr BEARDALL, William; Dr ZHAO, Yiren; Prof. STAN, Guy-Bart

Presenter: LI, Zehui

Session Classification: Contributed Talks

Track Classification: Contributed Talks