Developing an Inverse Reinforcement Learning Methodology to Predict the Progression of Colorectal Cancer

Authors: Silba S. Dowell[1], Daniel Hintz[1], Tyson Limato[1], Shad Sellers[1], Milana Wolff[1], Nicholas Chia[2], Liudmila S. Mainzer[1]
[1]University of Wyoming, Laramie WY [2]Argonne National Lab

A scalable LLM using IRL for predictive cancer research

Introduction

Biomedical research in cancer benefits from using high dimensional 'omics datasets, i.e., genomic, transcriptomic, proteomic, and phenotypic data. These are also large datasets due to the size of the human genome (~ 3 bln base pairs). As a result, computing requirements can be very high. Therefore, our team is developing a Machine Learning (ML) engine to identify causal mechanisms of an individual's cancer and predict its next stages of development in real-time without excessive need for expensive hardware.

Our ML engine is based on the Inverse Reinforcement Learning (IRL) algorithm because it closely parallels the step-by-step accumulation of somatic alterations in cancer evolution (Figure 1C). IRL is a specific form of learning from demonstration that attempts to estimate the reward function of a Markov decision process (MDP) from examples provided by the expert. For cancer, this reward function represents the possible sequences of mutations through which cancer evolves (Figure 1D), assigning transition probabilities for every mutation. Our machine learning engine uses high-performance computing in conjunction with several types of neural networks, such as LLMs, LSTMs, and DenseNets to reduce compute time and complexity.

Anticipated Results

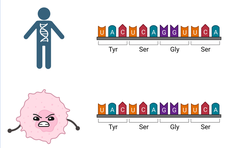
- Training on diverse patient data and including multiple features such as age, stage of cancer, lifestyle, etc., enables us to impart additional realism to the stochastic model of multiclonal phylogeny of tumors created by the LSTM state-action pair training.
- Using embeddings and dimensionality reduction, in addition to parallelization across multiple GPUS, results in faster compute time.
- Applying the same IRL pipeline across different infectious vectors, like viruses and many types of cancerous tumors, enables the extraction of meaningful disease features in each domain, indicating the generalizability of our approach.

Future Impact

- The use of IRL in reconstructing the evolutionary history of a tumor is a paradigm-shifting AI approach that can significantly improve our understanding of the molecular mechanisms of disease and accurately predict its course of progression, thus enabling effective drug target identification for treatment.
- This model prototype would be used for patient diagnosis and prognosis that may be used in hospitals nationwide.
- Our model functionality may be effective for the prediction of other types of cancers.

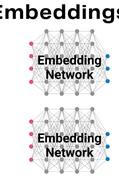
Methods

Input Data



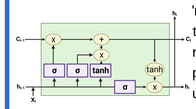
The input data is a FASTA file containing the patient's DNA sequences retrieved from Mayo Clinic's in-house patient data storage. In addition, we are also utilizing the COVID viral sequences from GISAID [7].

Genomic Embeddings



Due to the high order of dimensionality and length of the input data, an embedding network is used to encode the DNA sequence into a numerical representation to conserve computational resources and time. These numerical values are known as genomic embeddings. These embeddings encode both the state and actions for our IRL model.

LSTM for IRL state-action Pairs



A Long Short Term Memory model is a recurrent neural network that retains 'memory' of information. LSTM will use the state action pairs to learn the reward function and policy of the phylogenetic trajectories in order for us to produce all the possible trajectories that a hypothesized cancer cell may take during evolution.

Expert Demos & Reward Function

$\zeta_i = \{(s_{i,0}, a_{i,0}), (s_{i,1}, a_{i,1}), \ldots, (s_{i,T}, a_{i,T})\}$

The reward function represents the trajectories by which cancer evolves, assigning transition probabilities for every action. The likelihood of an evolutionary trajectory ζ, is given by the product of the likelihoods of each action along the trajectory, which are defined by the reward function R and exploration/exploitation hyperparameter ζ. In order to acquire the optimal trajectory, the model is trained on expert demonstrations.

We used PhyloWGS to find the VAFs for tumor subclones and the Augur pipeline to compute the phylogenetic tree for the cancer population and the mutations between each node. Finally, we used the mutations to create the state-action pairs for the expert demonstrations.

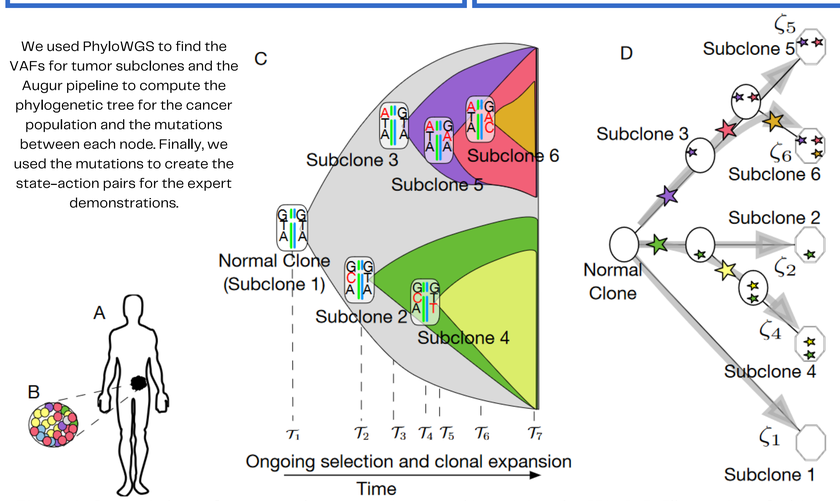


Figure 1. Genomic basis for expert demonstrations used in training the IRL. **A–B:** Illustration of a biopsy sample from a patient colon tumor. **C:** A model of clonal evolution is constructed based on the VAFs from inferred subclonal populations (refer to methods blob). **D:** The resultant phylogenetic tree is used to construct state-action pairs (genomic base + mutation) for expert demonstrations to train the IRL algorithm.

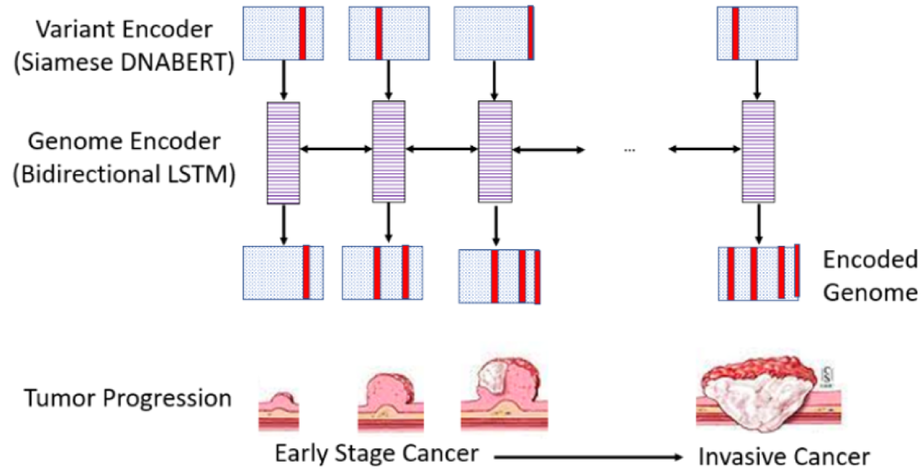


Figure 2. Siamese DNABERT and Bidirectional LSTM are both pretrained in correspondence with different tumor subclone samples.

Computational Theory

- Model & Tokenizer: DNABERT2 (Embedding Network)
- Optimizer: AdamW (Progressive LR: 3e-5)
- Scheduler: Cosine Annealing Restarts/
- Data pipeline: gradient check-pointed Distributed Data Parallel through HuggingFace Accelerate and high-level Python C extensions provided through the Nvidia CuDNN framework.
- Testing with Oracle single and multi-node Research Cloud compute resources for evaluation.

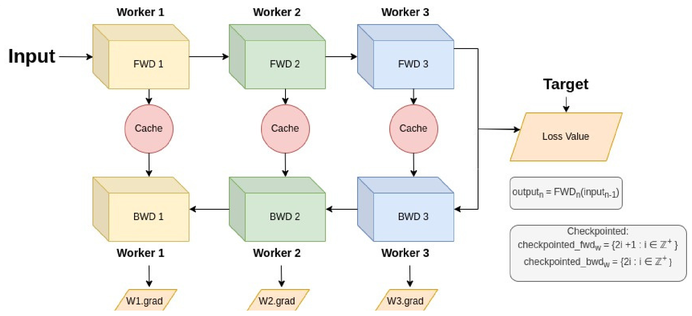


Figure 4. Gradient Checkpointed Distributed Data Parrallel

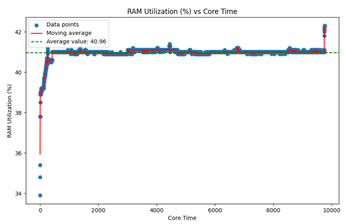


Figure 4. RAM usage for DNABERT

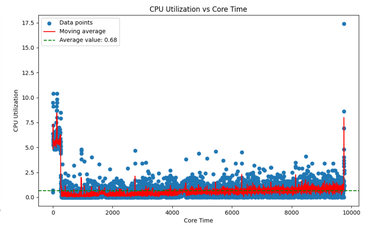


Figure 5. CPU usage for DNABERT

Discussion

Our work optimizes the compute aspects of a pioneering application in IRL for colorectal cancer [1]. This methodology, in combination with the use of genomic data embeddings, displays exceptional potential. Together, they will enable physicians to identify the causal mutations for a tumor in real-time with greater accuracy due to significant dimensionality reduction, keeping the compute time low. This is in contrast with approaches like Support Vector Machines (SVM), which escalate genomic input vectors into higher dimensional space to identify hyperplanes related to classes [5]. In contrast, prior ML models, such as decision trees, Bayesian networks, and SVMs [5] [6], classify cancer types with a lower resolution of tumor progression compared to our new methodology. Moreover, they do not permit reconstruction of the entire evolutionary path for multiclonal populations. Our work will result in a versatile and maintainable software for testing in clinical environments.

References

[1] John Kalantari, Heidi Nelson, and Nicholas Chia. The unreasonable effectiveness of inverse reinforcement learning in advancing cancer research. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 34, pages 437–445, 2020.
[2] Nicholas Chia, Carl R Woese, and Nigel Goldenfeld. A collective mechanism for phase variation in biofilms. Proceedings of the National Academy of Sciences, 105(38):14597–14602, 2008.
[3] Nigel Goldenfeld and Carl Woese. Biology's next revolution. Nature, 445(7126):369–369, 2007.
[4] Nigel Goldenfeld and Carl Woese. Life is physics: evolution as a collective phenomenon far from equilibrium. Annu. Rev. Condens. Matter Phys., 2(1):375–399, 2011.
[5] Konstantina Kourou, Themis P Exarchos, Konstantinos P Exarchos, Michalis V Karamouzis, and Dimitrios I Fotiadis. Machine learning applications in cancer prognosis and prediction. Computational and structural biotechnology journal, 13:8–17, 2015.
[6]Jiadong Chu, NA Sun, Wei Hu, Xuanli Chen, Nengjun Yi, and Yueping Shen. The application of bayesian methods in cancer prognosis and prediction. Cancer Genomics & Proteomics, 19 (1):1–11, 2022.
[7] Elbe, S. and Buckland-Merrett, G. (2017) Data, disease and diplomacy: GISAID's innovative contribution to global health. Global Challenges, 1:33-46. doi: 10.1002/gch2.1018 PMCID: 31565258
[8] Yanrong Ji, Zhihan Zhou, Han Liu, and Ramana Davulrui. DNABERT: pre-trained Bidirectional Encoder Representations from Transformers model for DNA-language in genome. Bioinformatics. 37(15):2112–2120, 2021.
[9] Zvyagin, Maxim, Alexander Brace, Kyle Hippe, Yuntian Deng, Bin Zhang, Cindy Orozco Bohorquez, Austin Clyde, et al. "GenSLMs: Genome-Scale Language Models Reveal SARS-CoV-2 Evolutionary Dynamics." Preprint. Bioinformatics, October 11, 2022. https://doi.org/10.1101/2022.10.10.511571.

Acknowledgments

Thanks to WY-INBRE, WY-REDD, and the University of Wyoming Department of Statistics for their financial support.