



**Vilnius
universitetas**

Doktorantas:
Brendonas Stakauskas
2022-2026

Darbo vadovas:
Dr. Virginijus Marcinkevičius

2023/2024 metai
II metai, II pusmetis



**Giliais neuroniniais tinklais grįstų
mašininio mokymo metodų taikymas
viruso mutacijų trajektorijai
prognozuoti**

TURINYS

1. Problemos apibrėžimas, tyrimo objektas ir tikslai
2. Studijų plano vykdymas
3. Trumpas per pusmetį gautų mokslinių rezultatų pristatymas
4. Kito pusmečio darbo planas



**Problemos
apibrėžimas,
tyrimo
objektas
ir tikslai**

Tyrimo objektas

Virusų baltymų sekos ir giliaisiais neuroniniais tinklais grįsti mašininio mokymo algoritmai skirti prognozuoti viruso mutacijas.

Tyrimo problemos

- Sąryšio tarp viruso proteinų sekų nustatymas;
- Duomenų aibės, atspindinčios istorines mutacijas, sudarymas;

Tyrimo tikslas

Sukurti giliaisiais neuroniniais tinklais ir natūralios kalbos apdorojimo algoritmais grįstą metodą leidžianti numatyti viruso mutacijos trajektoriją.

Tyrimo uždaviniai

- Atlikti literatūros analizę, išanalizuoti *state-of-the-art* algoritmus viruso baltymo mutacijų prognozavimui.
- Atkartoti literatūroje pateikiamų metodų rezultatus.
- Sukurti metodą duomenų aibės, atspindinčios istorines mutacijas, sudarymui.
- Sukurti duomenų aibę tyrimui.
- Pasiūlyti naują metodą viruso mutacijoms numatyti.
- Atlikti eksperimentinius tyrimus, palyginant pasiūlytą metodą su literatūroje aprašytais metodais.

Studijų plano vykdymas

Studijų planas, vykdymo suvestinė

Studijų metai	Egzaminai		Dalyvavimas konferencijose				Publikacijos				
			Nacionalinėse		Tarptautinėse		Su citavimo rodikliu		Be citavimo rodiklio		Buklė
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	
I (2022/2023)	2	2	1	0	0	1	0	0	1	1 (konferencijos darbų medžiagoje)	
II (2023/2024)	2	2	1	0	0	1	0	0	1	0	
III (2024/2025)	0	0	0	0	1	0	1	0	0	0	
IV (2025/2026)	0	0	0	0	1	0	1	0	0	0	

Ataskaitinio pusmečio darbo planas ir jo įvykdymas

Egzaminai 2023/2024 (II pusmetis)

Planas	Įvykdyta	Būklė
Natūralios kalbos apdorojimas	2024-06-27	Egzaminas <u>išlaikytas.</u>

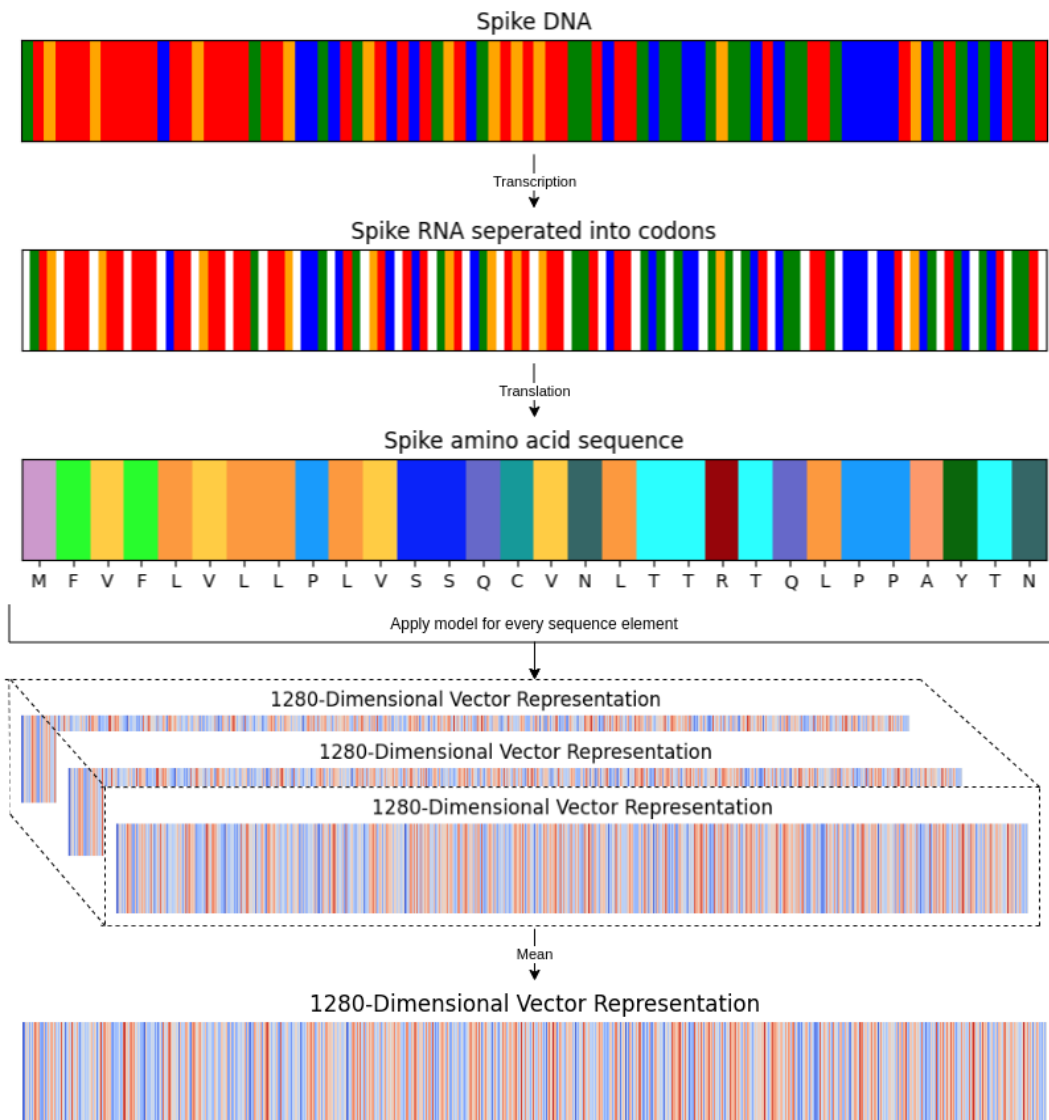
Visų mokslinių tyrimų ir disertacijos rengimo etapai

1. Mokslinių tyrimų disertacijos tema apžvalga ir analizė
2. Mokslinio tyrimo vykdymas
 1. Tyrimo metodikos sudarymas
 2. Teorinis tyrimas
 3. Empirinis tyrimas
 4. Gautų duomenų analizė, apibendrinimas, išvadų parengimas
3. Atskirų daktaro disertacijos dalių (tyrimo metodikos, rezultatų, ginamų teiginių, išvadų, ir kt.) parengimas
4. Daktaro disertacijos parengimas ir svarstymas padalinyje
5. Daktaro disertacijos gynimas

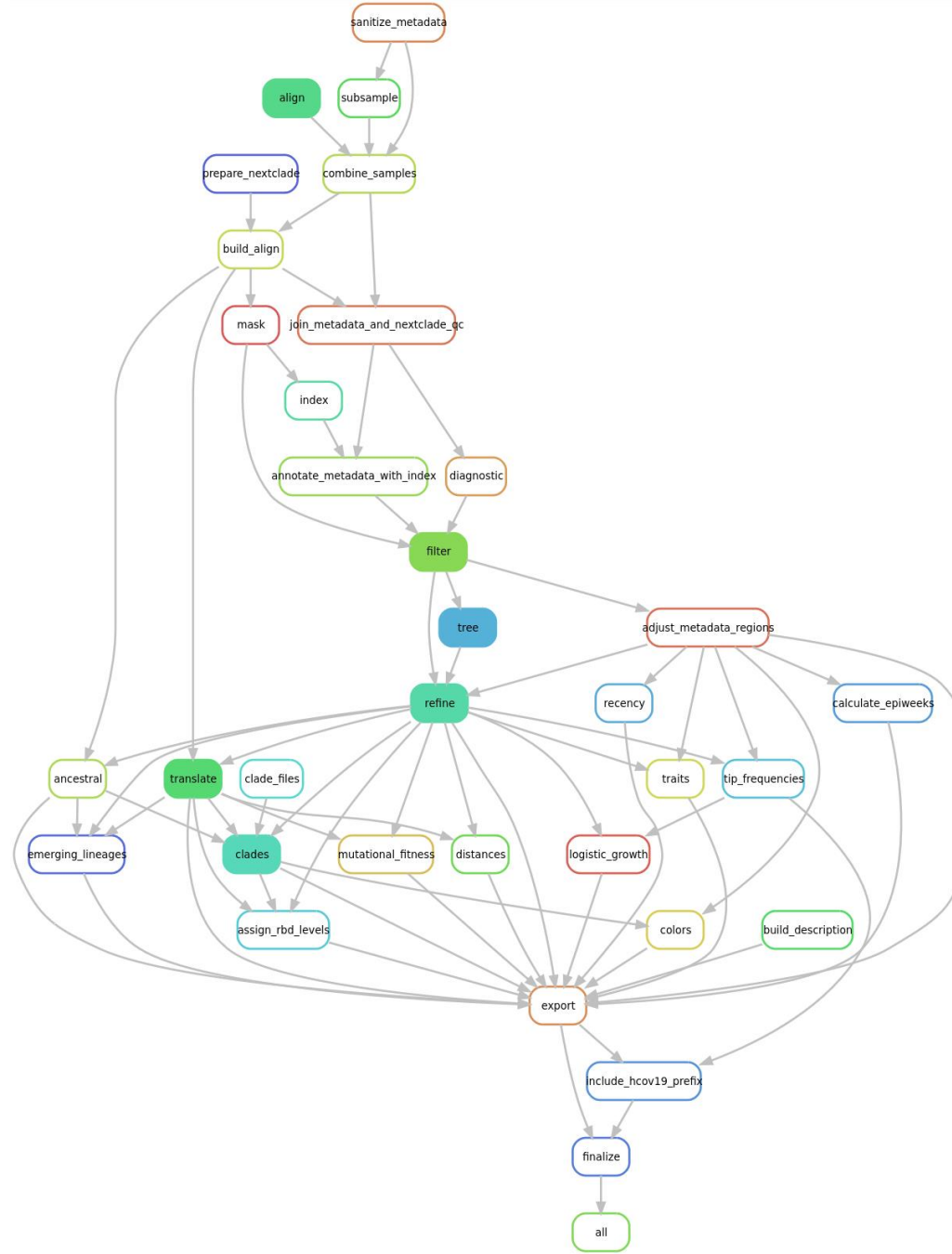
**Trumpas per
pusmetį gautų
mokslinių rezultatų
pristatymas**

Duomenų vizualizavimas

Naudojami ESM modelio įterpiniai siekiant atvaizduoti viruso spyglio amino rūgščių seką.



DNR transformacijos į įterpinį proceso iliustracija



Duomenų apdorojimo procesas pagal nextstrain

Duomenų rinkinys

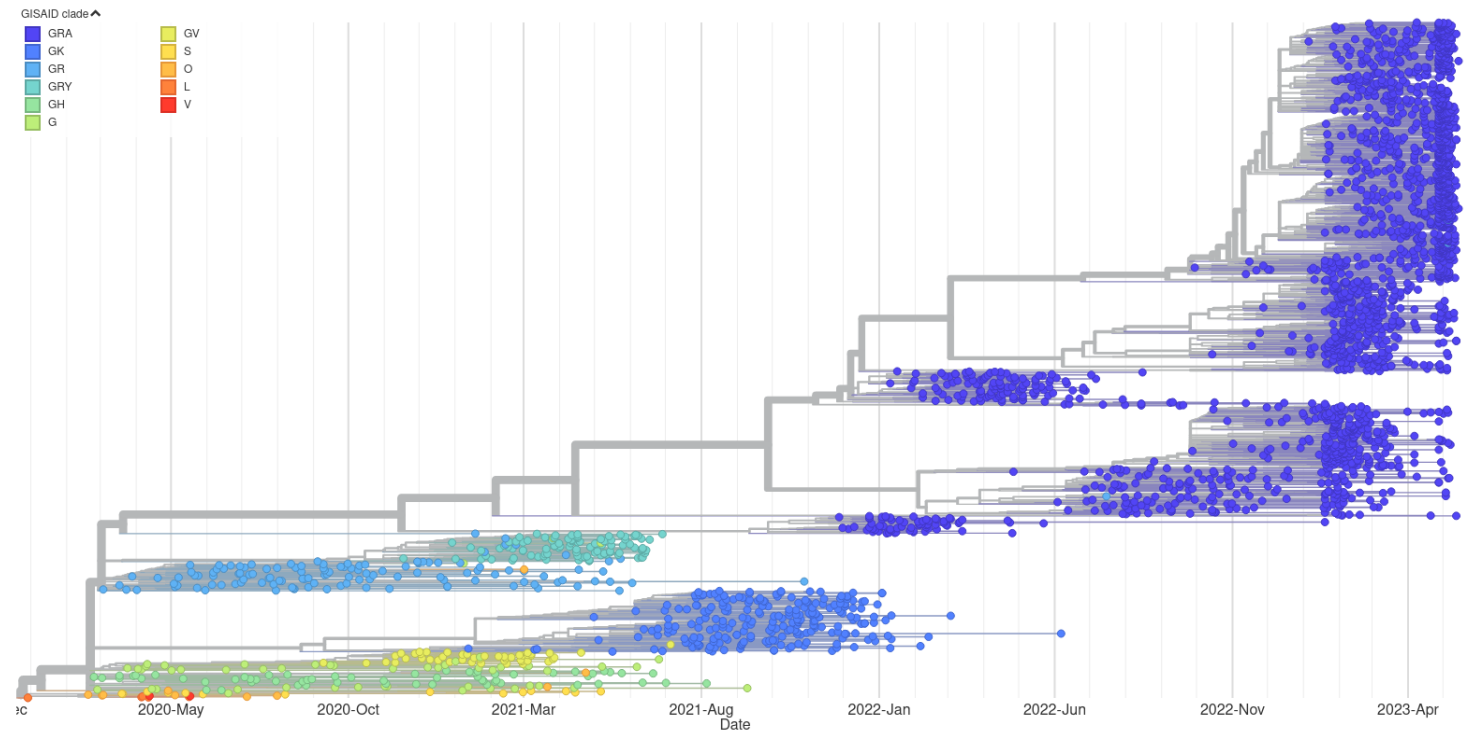
- 41387 viruso sekų
- Duomenys surinkti Lietuvoje
- 2020 sausis – 2023 kovas
- Filogenetiniam medžiui parinktas IQ-Tree algoritmas
- Spike genomas (1273 ilgio)

Modeliai

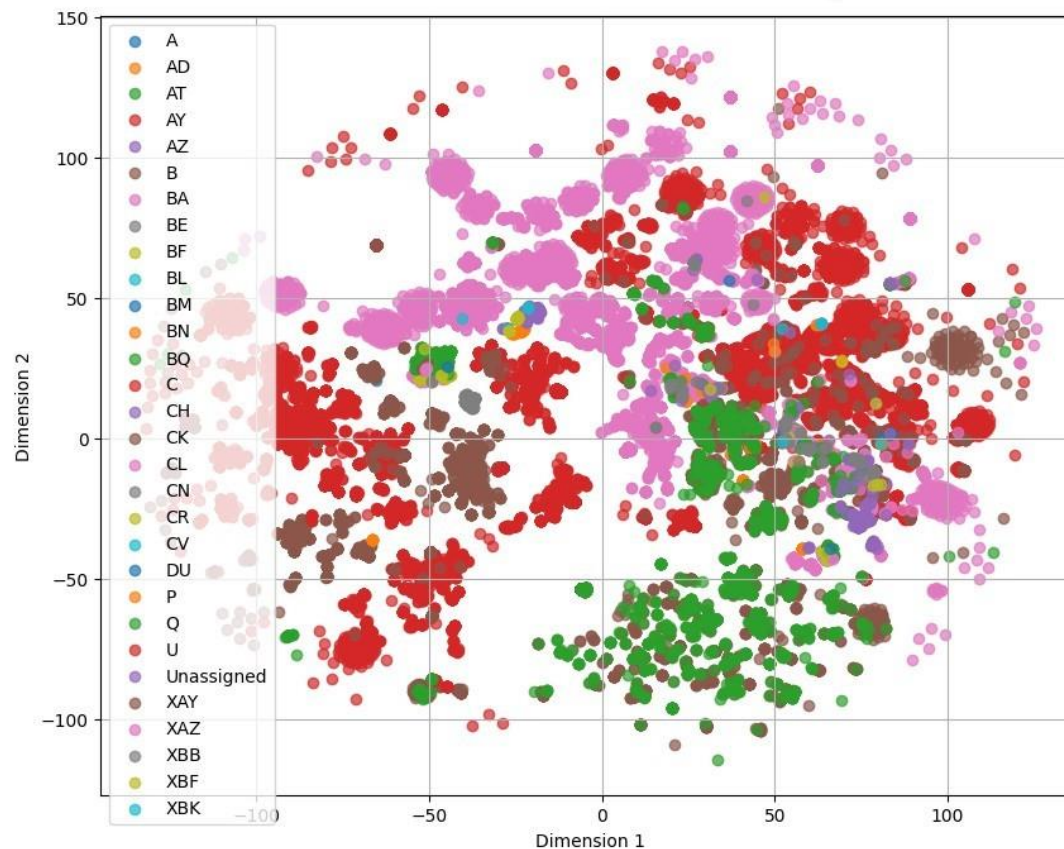
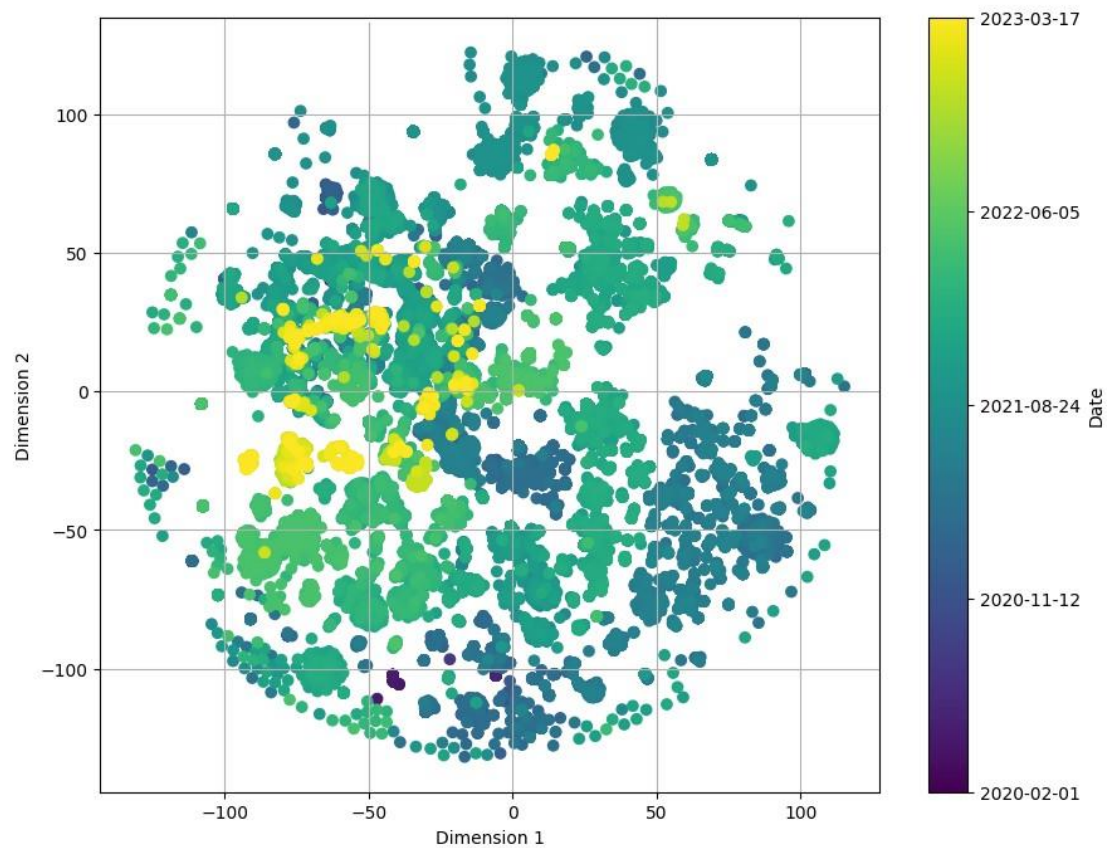
- ESM-1b
 - Parametrų skaičius -- 650 mln.
 - Išvestis -- 1280 reikšmių
- ESM-2
 - Parametrų skaičius -- 3 mlrd.
 - Išvestis -- 2560 reikšmių

Filogenetiniai medžiai

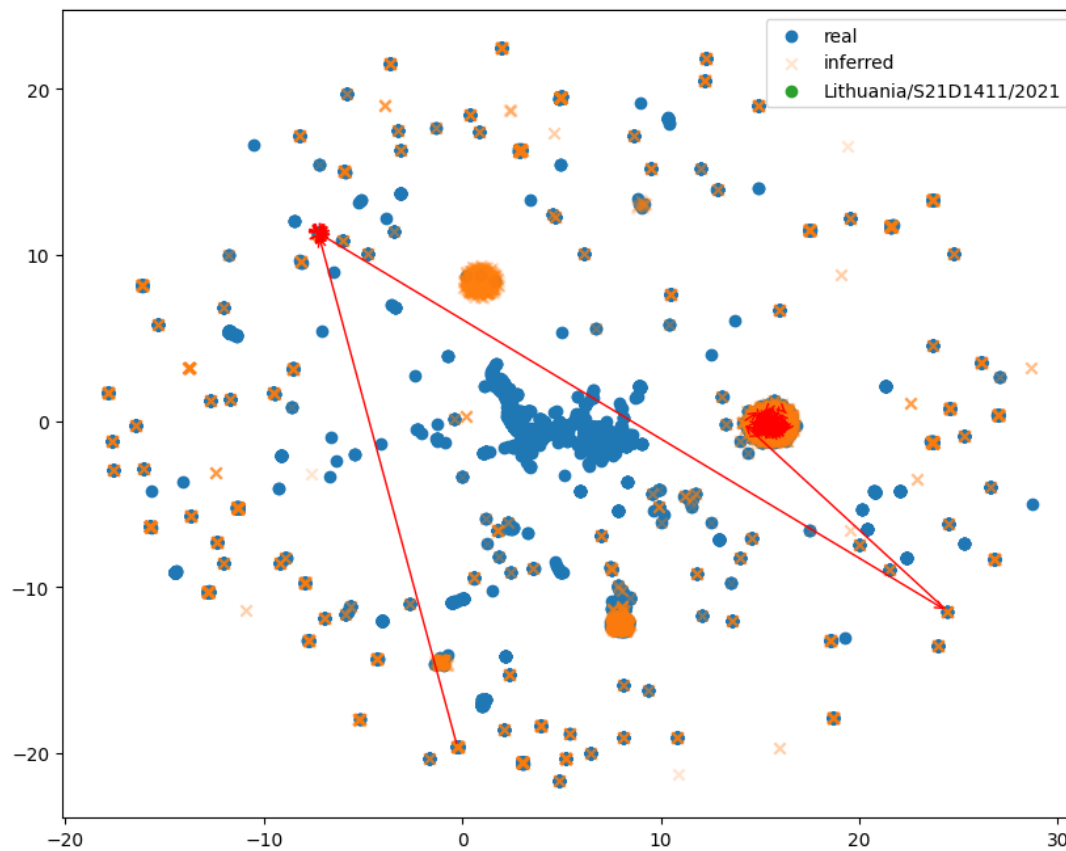
Filogenetinis medis, tai medis nurodantis evoliucinius ryšius tarp įvairių organizmų.



Duomenų vizualizavimas



Medžio vizualizavimas



Tarpinės išvados

- Dalis mutacijų nutinka pozicijose, kurios nebuvo įtrauktos į analizę.
- Galima imti visą viruso seką, bet egzistuoja rizika, kad duomenys supanašės, dėl vidurkinimo.
- Sudarytas medis gali būti per gilus, skirtumai tarp viršunių gali būti per maži norint stebėt kitimą.

Kito pusmečio darbo planas

Mokslinė veikla:

- Skirtingų algoritmų palyginimas ir tinkamiausių tolimesniems tyrimams atrinkimas.
- Atrinktų algoritmų įgyvendinimas, modifikavimas ar naujų algoritmų kūrimas, sprendžiant apibrėžtus uždavinius.
- Eksperimentinis sukurtų modifikacijų tyrimas, analizuojant jų efektyvumą.

Dalyvavimas konferencijose

Dalyvavimas konferencijoje Lietuvoje (DAMMS)

Publikacijų rengimas

Mokslinių tyrimų disertacijos tema apžvalga (konferencijos darbų medžiagoje)