



VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

FUNDAMENTINIŲ MOKSLŲ FAKULTETAS

CHEMIJOS IR BIOINŽINERIJOS KATEDRA

Taisija Dėmčėnko

**BALTYMO – LIGANDO JUNGIMOSI IN SILICO VERTINIMO FUNKCIJOS  
KŪRIMAS NAUDOJANT MAŠININĮ MOKYMĄSI**

**CREATING MACHINE LEARNING BASED SCORING FUNCTION FOR  
PROTEIN – LIGAND BINDING IN SILICO**

Baigiamasis bakalauro darbas

Bioinžinerijos studijų programa, valstybinis kodas 612J76001

Biotechnologijos studijų kryptis

Vilnius, 2020

# Turiny

<b>1</b>	<b>Įvadas . . . . .</b>	<b>3</b>
<b>2</b>	<b>Literatūros apžvalga . . . . .</b>	<b>3</b>
2.1	Baltymų sąveikos su skirtingomis molekulėmis . . . . .	3
2.1.1	Baltymo su ligandu surišimo konstanta . . . . .	3
2.1.2	Baltymo struktūros pasikeitimai skirtingose sąlygose . . . . .	3
2.2	Baltymo ir ligando sąveikos modeliavimas . . . . .	3
2.2.1	Ligando įvedimas į baltymą . . . . .	3
2.3	Vertinimo funkcijos baltymo sąveikos prognozei . . . . .	3
2.3.1	Klasikinės vertinimo funkcijos . . . . .	3
2.4	Mašininis mokymasis . . . . .	4
2.4.1	Mašininio mokymosi algoritmai . . . . .	4
2.4.2	Mašininio mokymosi taikymas bioinformatikoje . . . . .	4
	<b>Literatūra . . . . .</b>	<b>4</b>

# **1. Įvadas**

## **2. Literatūros apžvalga**

### **2.1. Baltymų sąveikos su skirtingomis molekulėmis**

Trumpai apie baltymą kaip molekulę Trumpai apie baltymo aktyvų centrą ir galinčius susikurti ryšius Trumpai apie sąveikas baltymo su ligandais, DNR, kitais baltymais Pavyzdys kiekvienai iš sąveikų 3-4 pavyzdžius sąveikos su ligandais (pasirinkti skirtingų tipų baltymus bei ligandus) Aprašyti skirtingus aktyvius centrus (su hemu ir pan.)

#### **2.1.1. Baltymo su ligandu surišimo konstanta**

Daugiau išsiplėsti. Kas vadinama ligandu? Sąveikos formulė, surišimo konstanta Būdai matuoti surišimo konstantą in vitro (ELISA ir t.t.)

#### **2.1.2. Baltymo struktūros pasikeitimai skirtingose sąlygose**

„Spynos ir rakto“, „indukuoto įtalpinimo“, „konformacijos išrinkimo“ teorijos Problema surišimo konstantos nustatymo (ką pasakojo Vytautas)

### **2.2. Baltymo ir ligando sąveikos modeliavimas**

Šiek tiek istorijos bei koks yra tame tikslas.

#### **2.2.1. Ligando įvedimas į baltymą**

„Docking“. Aprašyti kelius algoritmus.

### **2.3. Vertinimo funkcijos baltymo sąveikos prognozei**

Šiek tiek daugiau istorijos.

#### **2.3.1. Klasikinės vertinimo funkcijos**

Aprašyti 3-5 klasikinių vertinimo funkcijų algoritmus.

## 2.4. Mašininis mokymasis

Bendra atsiradimo istorija. Apmokymo ir testavimo principas. Regresija, klasifikacija: skirtumai (trumpai). Mašininio mokymosi taikymas.

### 2.4.1. Mašininio mokymosi algoritmai

Pradėti nuo pačio paprasčiausio Decision Tree, pasistengti įsigilinti į formules. Atsitiktiniai miškai. Atraminių vektorių mašinos.

### Dirbtiniai neuroniniai tinklai

Lašelis istorijos, taikymas. Skirtingi dirbtinio neuroninio tinklo tipai (4-6 pagrindinių) Konvoliuciniai neuroniniai tinklai – plačiau (nes šis bus taikomas darbe)

### 2.4.2. Mašininio mokymosi taikymas bioinformatikoje

Sujungti biologiją su mašininio mokymosi. Aprašyti 2-3 mašininio mokymosi algoritmus, taikomus bioinformatikos srityje. Aprašyti 3-4 mašininio mokymosi vertinimo funkcijas (paskutinė būtų referentinė: DeepVS[1]).

### Duomenų paruošimas pateikimui į modelį

Priminti apie duomenų kaupimo greitį ir apie tai, kad mašininio mokymosi efektyvumas stipriai priklauso nuo duomenų apimties. Aprašyti viešai prieinamas duombazes. Akcentuojant konvoliucinius tinklus, aprašyti duomenų virsmą į skaičius, kuriuos suvalgys mašininio mokymosi modelis. Priminti apie surišimo konstantų problemą.

### Vertinimo funkcijos įverčiai

Įverčiai, taikomi modelį įvertinti (AUC, confusion matrix ir kiti). Pateikti kelių mašininio mokymosi vertinimo funkcijų įverčio pavyzdžius (iš tų funkcijų aprašų). Tarpusavyje jas palyginti nelabai išeis, nes reikalingi vienodi duomenys.

## Literatūra

- [1] Janaina Cruz Pereira, Ernesto Raúl Caffarena, and Cicero Nogueira Dos Santos. Boosting Docking-Based Virtual Screening with Deep Learning. *Journal of Chemical Information and Modeling*, 56(12):2495–2506, 2016.