



**UFRJ**



CONCURSO PÚBLICO PARA PROVIMENTO EFETIVO DE VAGAS NO CARGO DE PROFESSOR DA CARREIRA DE MAGISTÉRIO SUPERIOR EDITAL Nº 54 DE 30 DE JANEIRO DE 2024 - OPÇÃO DE VAGA RP-001 - CENTRO DE CIÊNCIAS DA SAÚDE - CCS - INSTITUTO DE CIÊNCIAS BIOMÉDICAS - PROGRAMA BIOLOGIA CELULAR E DO DESENVOLVIMENTO - BIOINFORMÁTICA - PROFESSOR ADJUNTO A - 40H(DE)

## PROVA PRÁTICA

CANDIDATO: DR. GABRIEL RODRIGUES COUTINHO PEREIRA  
INSCRIÇÃO: 1701857



# Universidade

```
graph LR; U[Universidade] --- E[I. Ensino]; U --- P[II. Pesquisa]; U --- Ext[III. Extensão]; U --- Ad[IV. Administrativa];
```

I. Ensino

II. Pesquisa

III. Extensão

IV. Administrativa

## GRADUAÇÃO

### **Plano**

1. Atualizar ementas e planos de ensino;
2. Promover uma formação teórica e prática sólida;
3. Promover a capacitação técnica.

# I. ENSINO: GRADUAÇÃO

- I. Atualizar ementas e planos de ensino
  - Participar do Núcleo Docente Estruturante (NDE);
  - Promover a atualização de ementas e planos de ensino;
  - Incorporar tendências da **academia** e do **mercado**;
  - Disciplinas obrigatórias e proposição de eletiva(s).



# I. ENSINO: GRADUAÇÃO

- Contribuir com disciplinas obrigatórias e proposição de disciplina(s) eletiva(s).



## Contribuir em disciplinas existentes:

- e.g., Biologia Molecular de Célula.

} Experiência como tutor (UNIRIO)

## Proposição de novas disciplinas (teórico-práticas):

1. Bioinformática (60h);
2. Ciência de Dados para Biomedicina (60h).

# I. ENSINO: GRADUAÇÃO

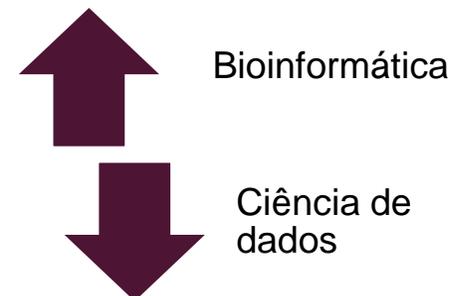
## Proposição de novas disciplinas (teórico-práticas):

### 1. Bioinformática (60h):

- Histórico e aplicações da Bioinformática;
- Estrutura física e organização gênica;
- Conceitos fundamentais em genômica;
- Introdução ao Python para Bioinformática;
- Bancos de Dados Biológicos;
- Alinhamento de Sequências;
- Padrões e Motivos Estruturais em Proteínas;
- Montagem de sequências de DNA;
- Análise Filogenética;

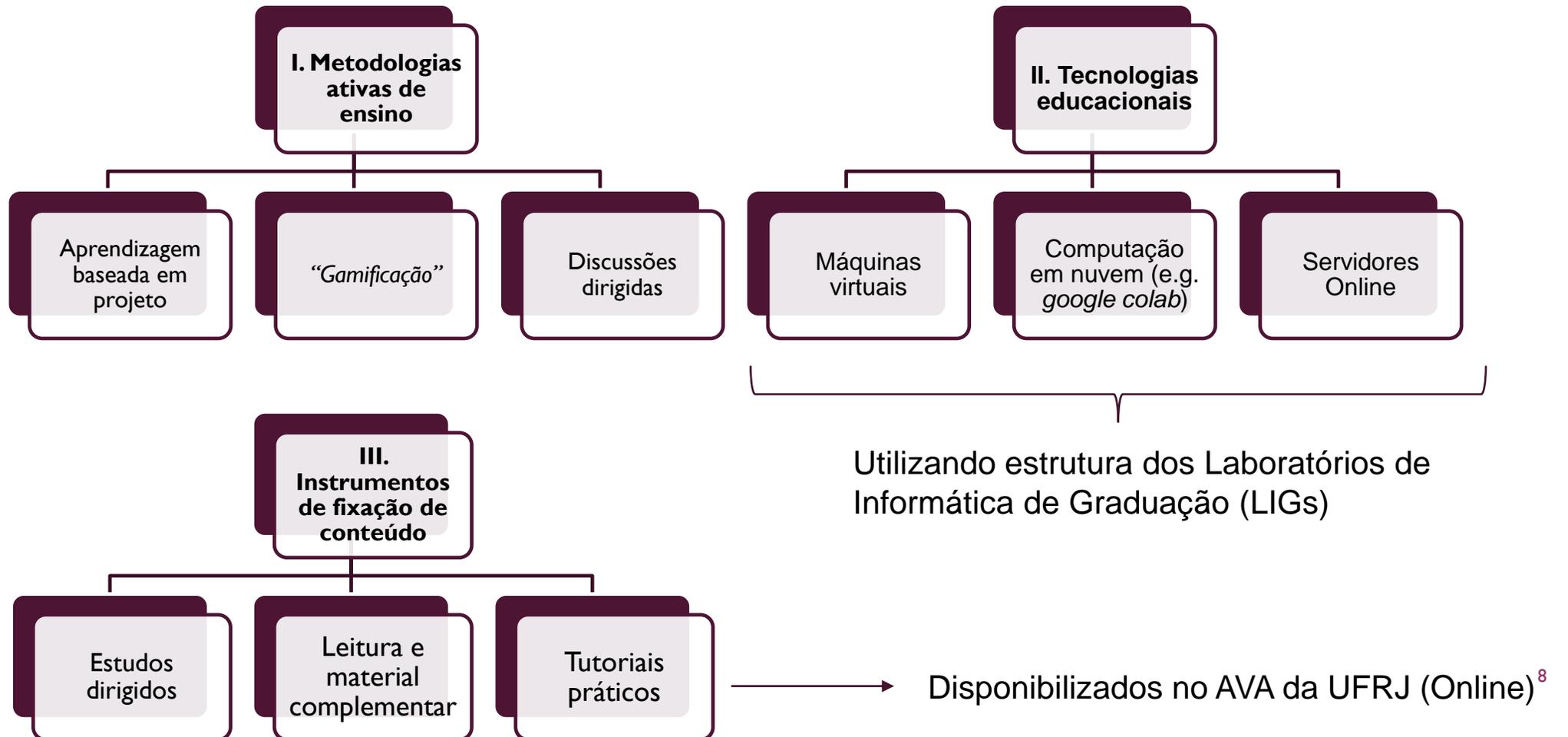
### 2. Ciência de Dados para Biomedicina (60h):

- Introdução ao Python para ciência de dados;
- Análise exploratória de dados;
- Visualização de dados;
- *Storytelling* com dados;
- LGPD e ética em dados;
- Aprendizado de máquinas;
- Otimização de modelos;
- Introdução ao *Business Intelligence*.

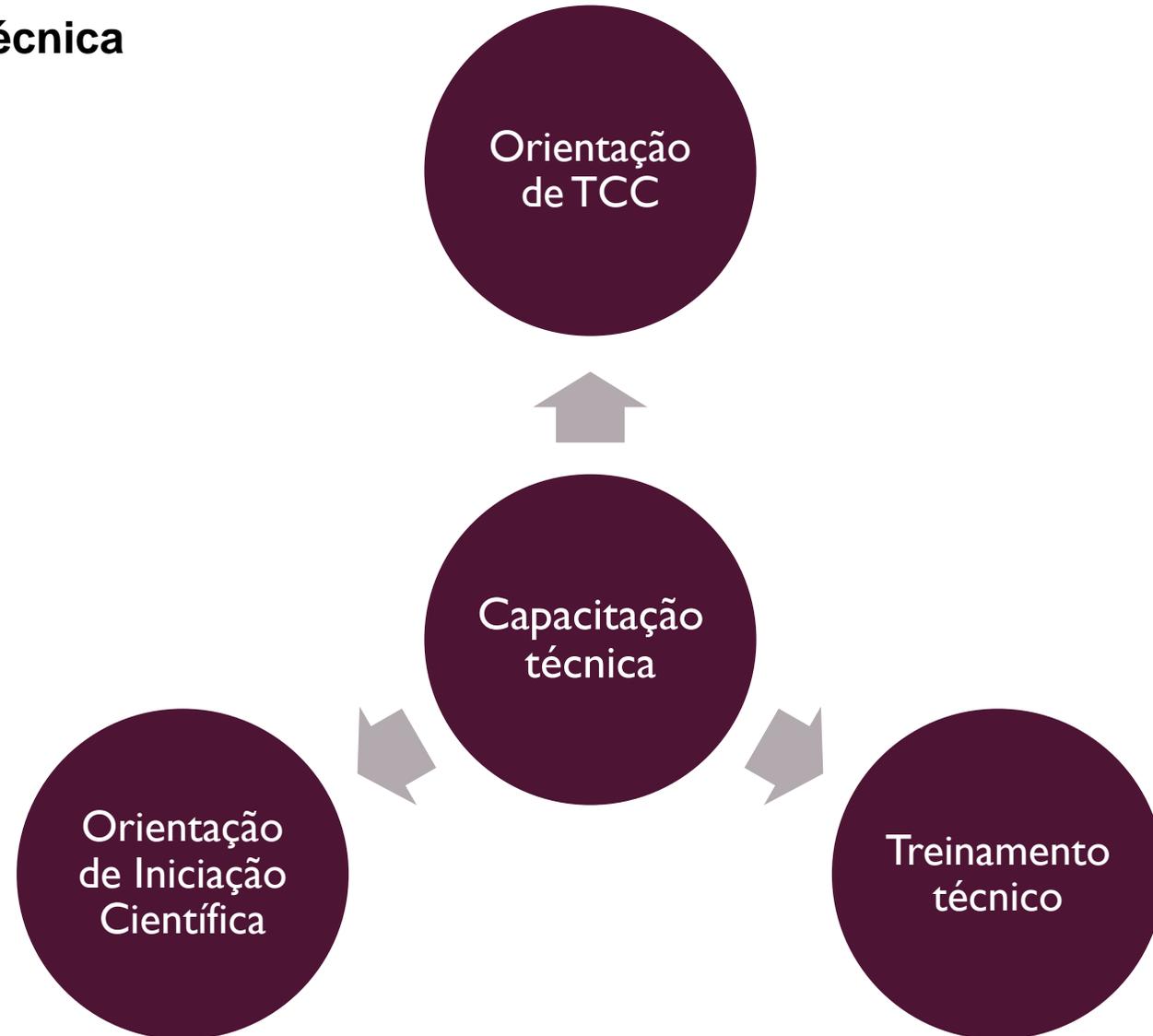


# I. ENSINO: GRADUAÇÃO

## 2. Promover uma formação teórica e prática sólida



## 3. Promover a capacitação técnica



## PÓS-GRADUAÇÃO

### **Plano**

1. Atuação em programas de Pós-Graduação;
2. Formação de pesquisadores;

# I. ENSINO: PÓS-GRADUAÇÃO

## 3. Contribuir com programas de Pós-Graduação:

- Submeter candidatura como membro permanente
- Contribuir com disciplinas já existentes;
- Proposição de disciplina eletiva.



**PPGBMC**  
PROGRAMA DE PÓS-GRADUAÇÃO  
EM BIOLOGIA MOLECULAR E CELULAR



Já atuo como  
coorientador



**Maria Letícia C. Barbosa**



Já atuo como<sup>11</sup>  
colaborador

# I. ENSINO: PÓS-GRADUAÇÃO

- Contribuir com disciplinas já existentes e proposição de disciplinas eletivas.

Disciplinas obrigatórias



Disciplinas eletivas



**Contribuir em disciplinas existentes:**

- e.g., Metodologia da Pesquisa e Química medicinal.

**Proposição de uma nova disciplina (teórico-prática):**

1. Modelagem molecular no planejamento de fármacos (45h).

# I. ENSINO: PÓS-GRADUAÇÃO

## Proposição da disciplina de modelagem molecular no planejamento de fármacos (45h).

### 1. Ementa:

- Banco de dados de proteínas e compostos químicos;
- Modelagem estrutural de proteínas;
- Mapa farmacofórico;
- *Docking* molecular;
- Fundamentos de modelagem QSAR;
- Farmacocinética e toxicologia *in silico*;
- *Docking* molecular;
- Triagem Virtual;
- Simulações de Dinâmica Molecular.

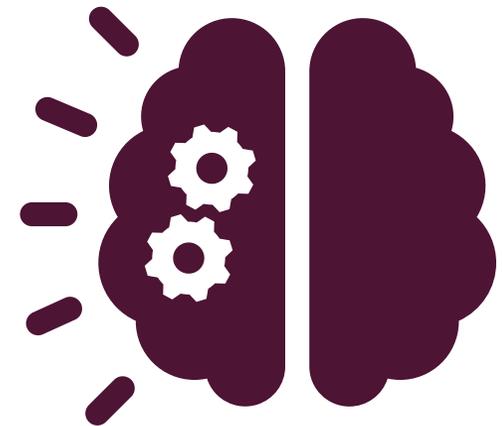
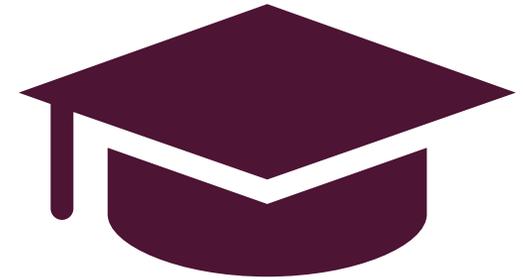
### Atividades práticas

- Uso de servidores online (e.g. DockThor, Pharmit);
- Computação em nuvem (Google Colab);
- Notebooks (Jupyter, Google Colab).

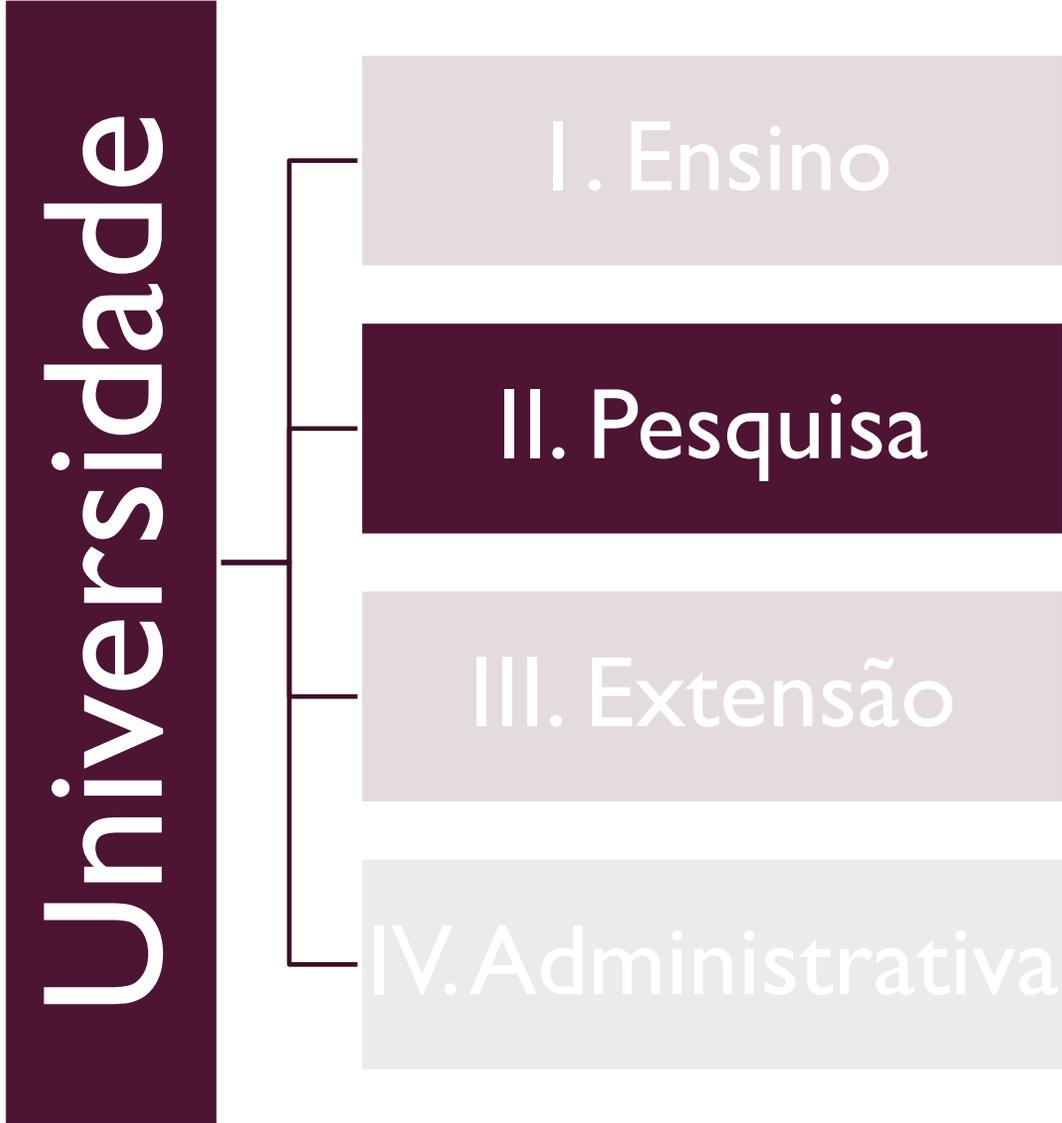
# I. ENSINO: PÓS-GRADUAÇÃO

## ■ 2. Formar pesquisadores:

- Treinamento técnico;
- Estimular a autonomia e o pensamento crítico;
- Fomentar a escrita científica;
- Publicação ética e responsável.



# Universidade



I. Ensino

II. Pesquisa

III. Extensão

IV. Administrativa

# II. PESQUISA

---

## **Plano**

1. Projeto de pesquisa;
2. Estabelecer e consolidar linhas de pesquisa na UFRJ;
3. Promover colaborações;
4. Captar fomento;
5. Publicação de artigos e participação em congresso.

# II. PESQUISA

## 1. Projeto de pesquisa

### Título:

- Desenvolvimento de uma Plataforma Interativa para a Predição do Impacto Funcional de Mutações *Missense*

### Justificativa:

- Milhões de novas variantes de nucleotídeo único (SNVs) são descobertas a cada ano;
- Caracterização do efeito funcional de todas essas mutações experimentalmente é inviável;
- Apenas 2% de todas as variantes já foram caracterizadas;
- Necessidade de ferramentas para filtrar e priorizar mutações a serem estudadas;
- Métodos disponíveis desatualizados, descontinuados/sem suporte, mal otimizadas, ou de difícil uso.

# II. PESQUISA

---

## 1. Projeto de pesquisa

### Objetivos

- Desenvolver, validar e otimizar um algoritmo competitivo de predição funcional para classificar novas mutações *missense*;
- Desenvolver um *software* que permita o uso do algoritmo por meio de uma interface gráfica amigável.

### Oportunidades

- Cenário de desenvolvimento de *software* em bioinformática ainda restrito no Brasil;
- Promover inovação e contribuir para o campo da bioinformática;
- Ferramenta com potencial didático.

### Inspiração

- LNCC: e.g. DockThor;
- UFMG: e.g. nAPOLI.

# II. PESQUISA

---

## Metodologia:



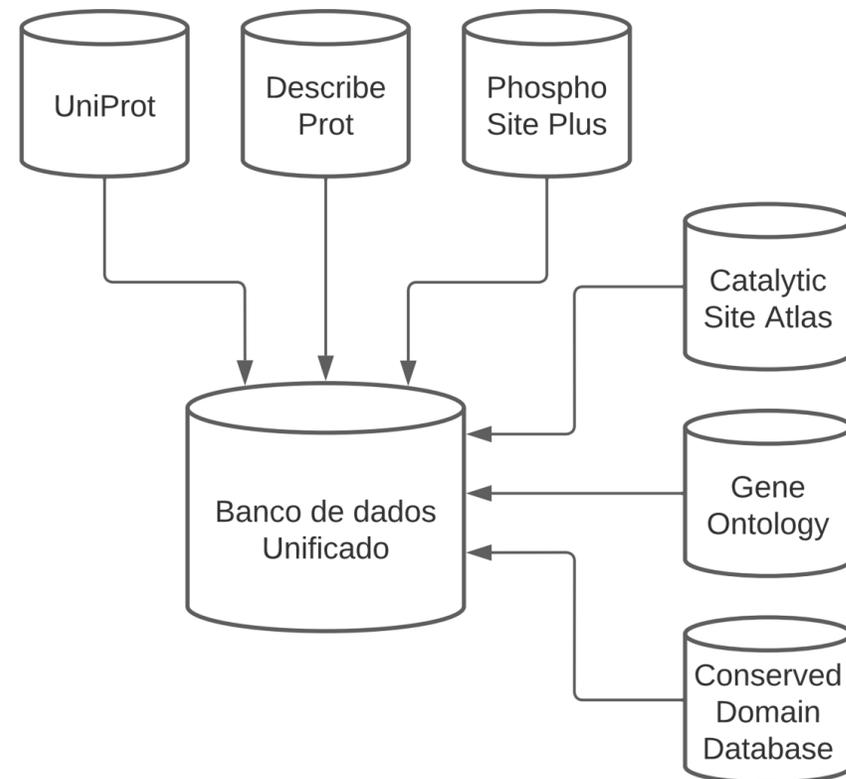
# II. PESQUISA

## 1. CONSTRUÇÃO DO BANCO DE DADOS UNIFICADO

### ▪ Obtenção de variáveis preditoras:

- UniProt;
- DescribeProt;
- PhosphoSite Plus;
- Catalytic Site Atlas;
- Gene Ontology;
- Conserved Domain Database;

### ▪ Integração de dados.



# II. PESQUISA

## 2. CONSTRUÇÃO DAS MATRIZES DE SUBSTITUIÇÃO

### ▪ Construção de matrizes de substituição 20x20:

- BLOSUM62;
- PSSM;
- Matrizes para características físico-químicas;
- Matrizes de frequência;
- Matrizes com alterações de energia livre.

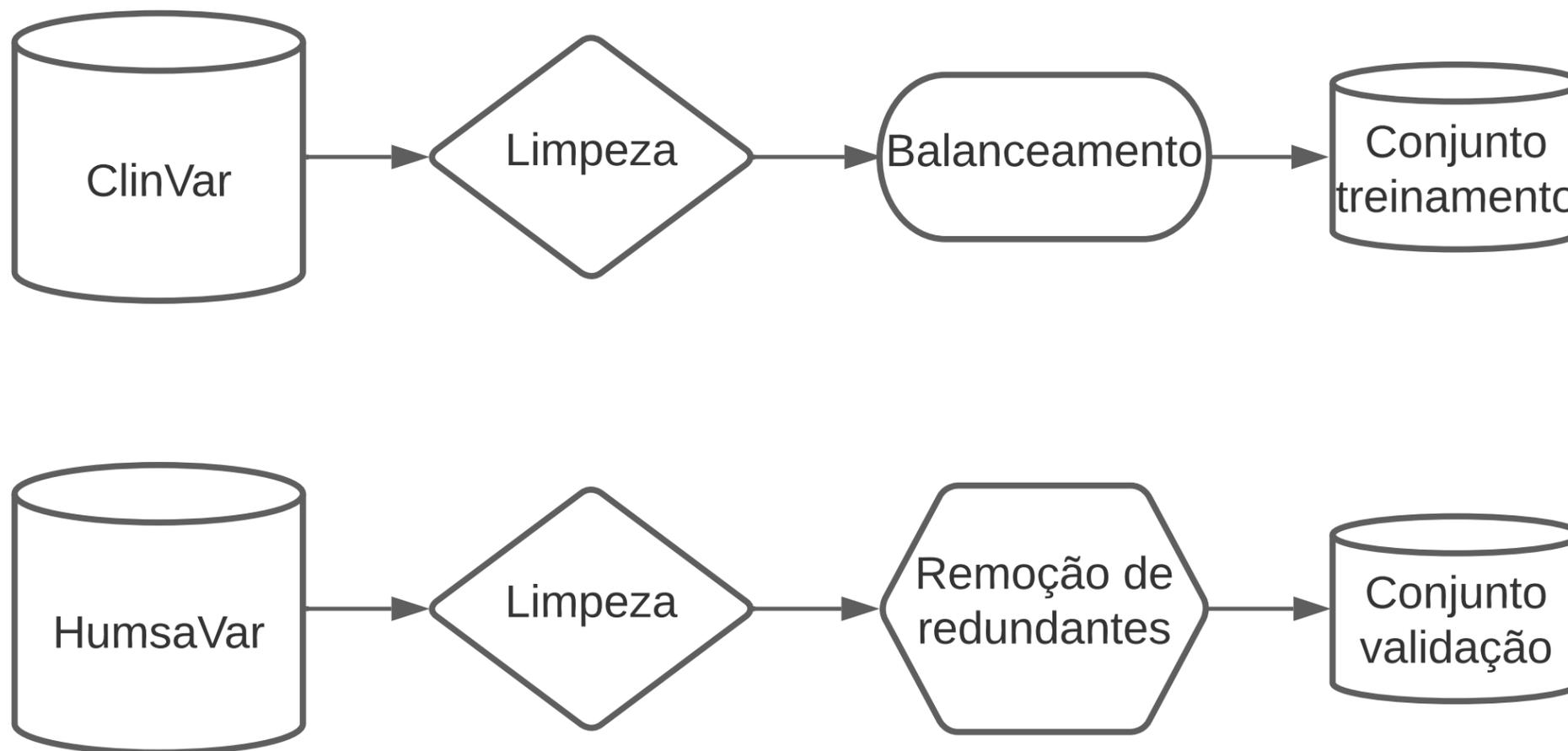
$$\begin{bmatrix} E_{1 \rightarrow 1} & E_{1 \rightarrow 2} & \dots & E_{1 \rightarrow j} & \dots & E_{1 \rightarrow 20} \\ E_{2 \rightarrow 1} & E_{2 \rightarrow 2} & \dots & E_{2 \rightarrow j} & \dots & E_{2 \rightarrow 20} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ E_{i \rightarrow 1} & E_{i \rightarrow 2} & \dots & E_{i \rightarrow j} & \dots & E_{i \rightarrow 20} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ E_{L \rightarrow 1} & E_{L \rightarrow 2} & \dots & E_{L \rightarrow j} & \dots & E_{L \rightarrow 20} \end{bmatrix}$$

20i x 20j

Matrizes de substituição

## II. PESQUISA

### 3. PREPARAÇÃO DO CONJUNTO DE TREINAMENTO E VALIDAÇÃO



# II. PESQUISA

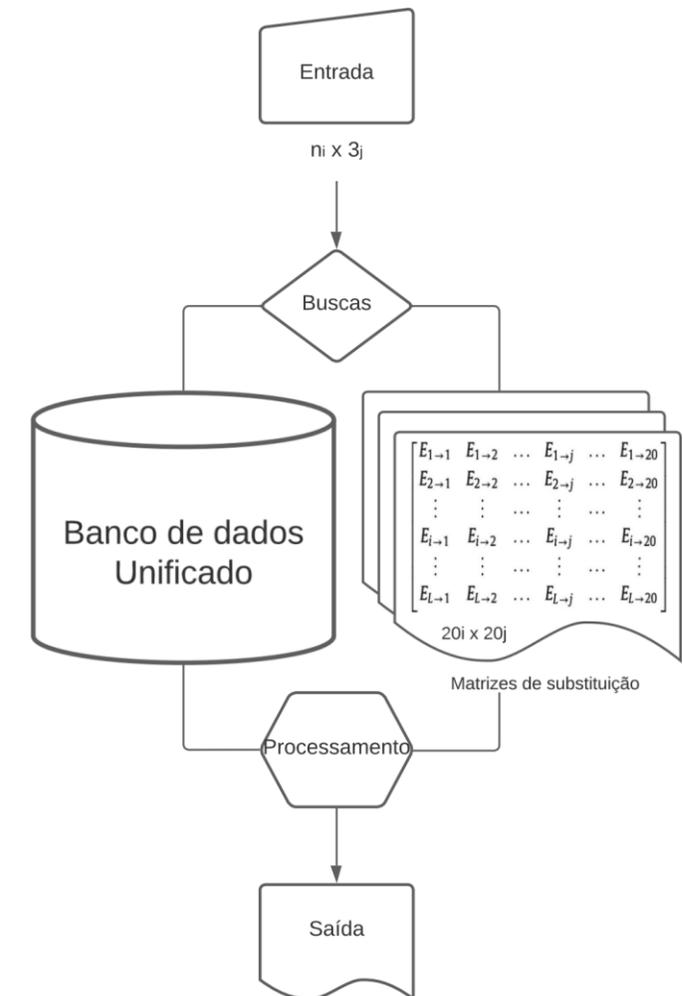
## 4. EXTRAÇÃO AUTOMATIZADA DE VARIÁVEIS PREDITORAS

### ▪ Criação de uma função em Python:

- Biblioteca Pandas;
- Baseada em iterações;
- Realiza buscas por variáveis preditoras correspondentes;
- Arquivo de saída com variáveis preditoras.

### ▪ Submissão dos conjuntos de treinamento e validação:

- Obtenção das variáveis preditoras correspondentes;
- Geração de arquivos de entrada para a etapa de modelagem e validação.



# II. PESQUISA

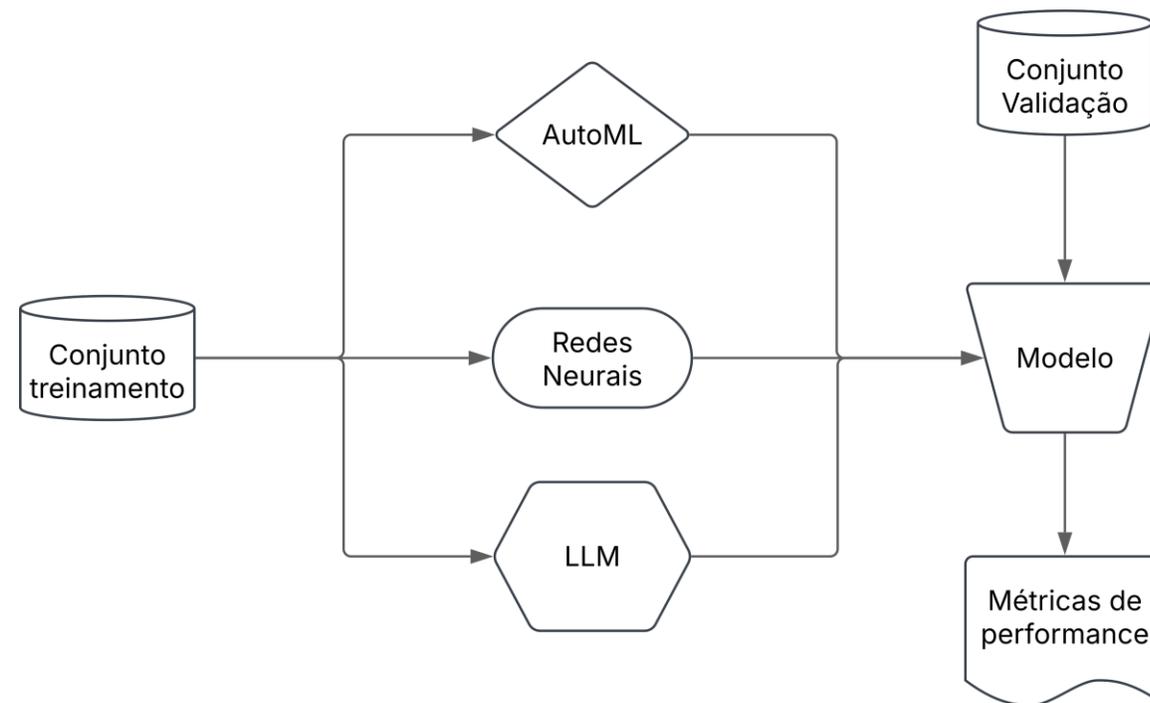
## 5. TREINAMENTO E SELEÇÃO DO MELHOR MODELO

### ▪ Criação de diferentes modelos:

- Conjunto de treinamento;
- AutoML na biblioteca Pycaret;
- Redes Neurais na biblioteca Keras-Tensorflow;
- ❖ Fine-tune VariPred (*LLM*).

### ▪ Validação dos modelos:

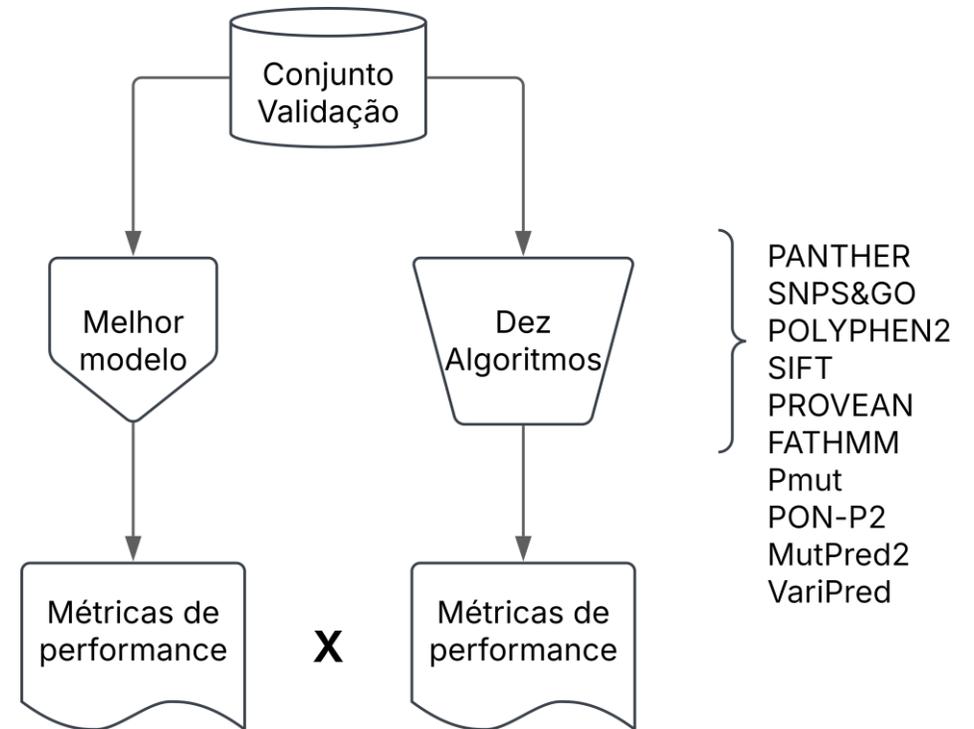
- Conjunto de validação;
- Acurácia, AUC-ROC, *recall*, precisão, *F1-score* e *MCC*;
- Escolha da melhor classe de modelos.



## 6. ANÁLISE DE VIABILIDADE DO MODELO

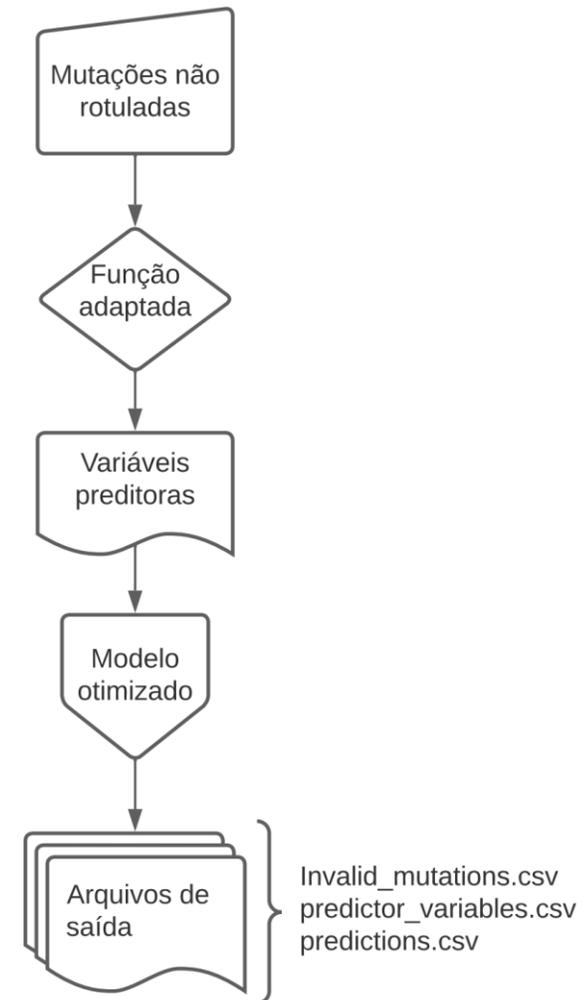
### ▪ Comparação de métricas de performance:

- Predições no conjunto de validação;
- Dez algoritmos ativos e disponíveis na literatura;
- Acurácia, AUC-ROC, *recall*, precisão, *F1-score* e *MCC*;



## 7. Construção de um pipeline automatizado

- **Otimizações da função para extração de variáveis:**
  - Divisão do banco de dados unificado por tipo de objeto;
  - Busca utilizando funções *map* e *lambda*.
- **Modelo final:**
  - Chamado.
- **Funcionamento do pipeline.**

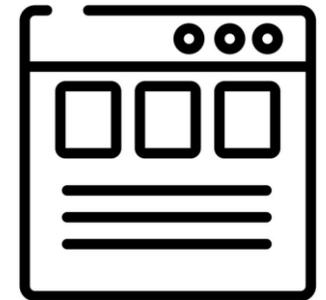


# II. PESQUISA

## 8. Desenvolvimento da plataforma gráfica

- **Software:**

- Utilizar o pipeline;
- Interface gráfica (GUI) – biblioteca PySimpleGui;
- Executáveis para Windows e Linux – biblioteca PyInstaller.

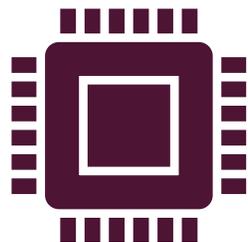


## RESULTADOS ESPERADOS

- Precisão equiparável ou superior a outros métodos disponíveis;
- Eficiência computacional;
- Interface gráfica amigável;
- Plataforma *no-code*;
- Utilização ampla;
- Ferramenta para fins didáticos.

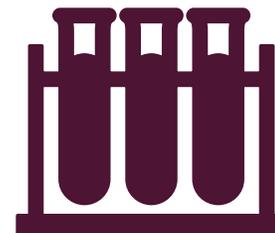
# II. PESQUISA

## 2. Estabelecer e consolidar linhas de pesquisa na UFRJ



### Estabelecer uma nova linha de pesquisa:

Desenvolvimento de *softwares* preditivos para bioinformática;  
Projeto proposto e potenciais desdobramentos.

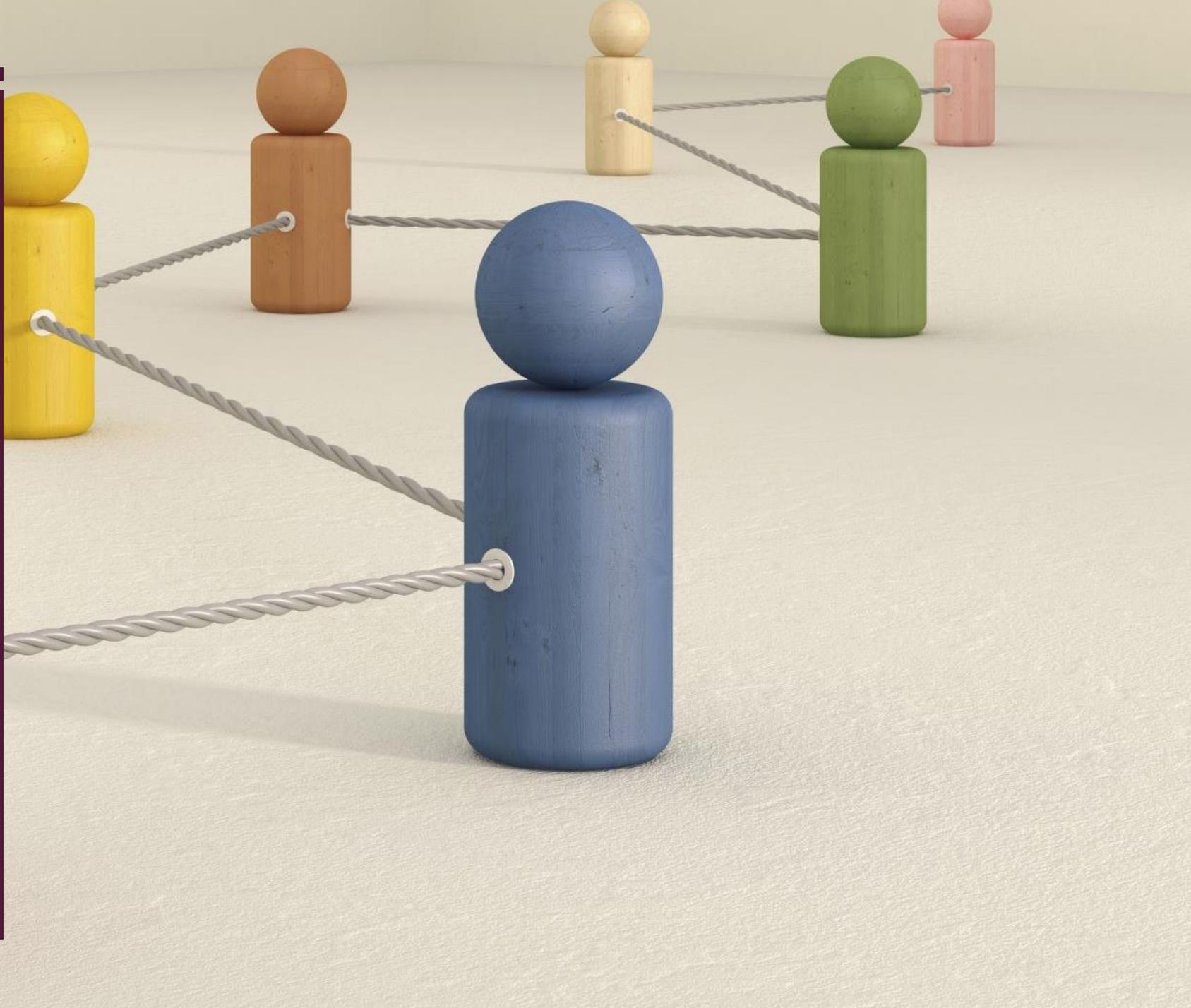


### Consolidar a linha de pesquisa já existente:

Planejamento de fármacos.  
Planejamento, síntese e avaliação biológica.

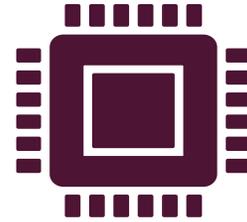


### 3. PROMOVER COLABORAÇÕES



# II. PESQUISA

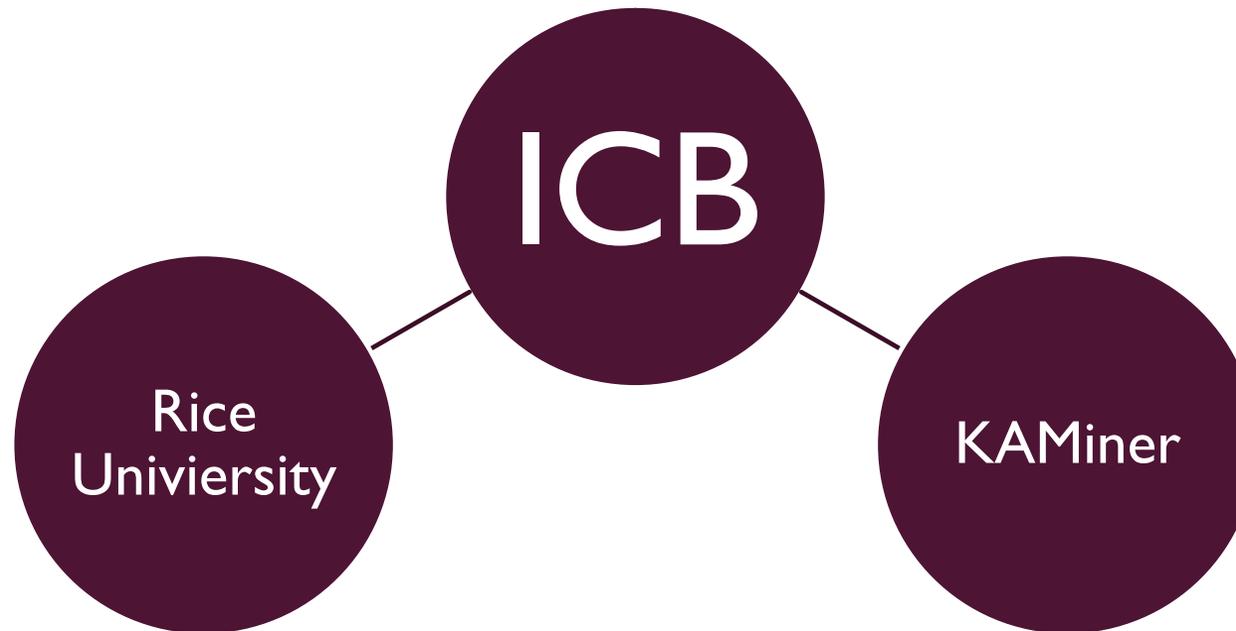
## 3. Colaborações:



Desenvolvimento de *softwares* preditivos para bioinformática



Lydia Kavraki



Altobelli Mantuan

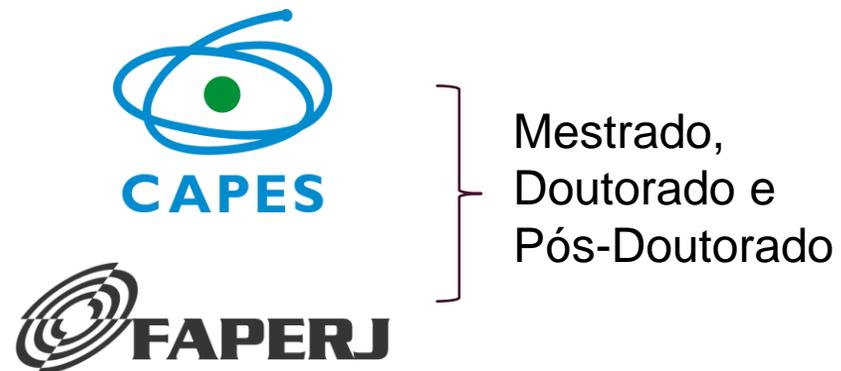
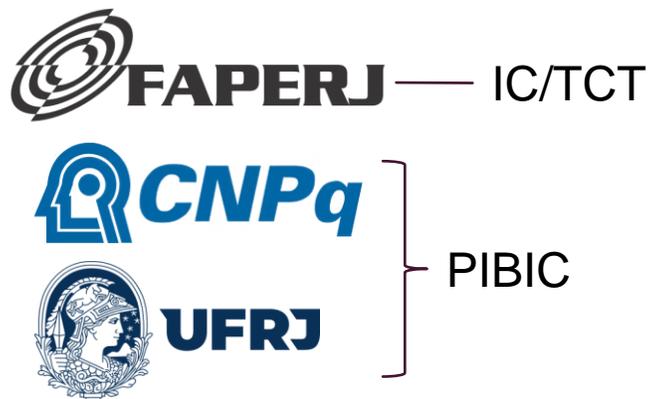


# II. PESQUISA

## 4. Captar fomento para a instituição;

- Contribuir para a infraestrutura institucional:
  - Equipamentos;
  - Obras;
  - e.g. Universal (CNPq), Auxílio a Pesquisa APQ1 (FAPERJ), Jovem Pesquisador Fluminense (FAPERJ), Jovem Cientista do Nosso Estado (FAPERJ).

- Captação de bolsas de fomento:



## 5. PUBLICAÇÃO DE ARTIGOS E PARTICIPAÇÃO EM CONGRESSO.



# Universidade

I. Ensino

II. Pesquisa

III. Extensão

IV. Administrativa



# LEARNING DROPS



saúde:

## A INTELIGÊNCIA ARTIFICIAL (IA) E O DIAGNOSTICO DE DOENÇAS

Pesquisadores desenvolveram um modelo treinado de forma semelhante ao chatGPT

Arraste pro lado →



@learningindrops



learningindrops

Seguindo ▾

Enviar mensagem



820 publicações

2.583 seguidores

6.785 seguindo

**Learning in Drops**

Educação

O LEARNINGINDROPS é um site educacional não formal com objetivo de divulgar conhecimento de modo prático e dinâmico... mais

🔗 [linkme.bio/learningindrops](https://linkme.bio/learningindrops)

<https://www.instagram.com/learningindrops/>



**UFRJ**

**course**ra

**Plataforma de cursos gratuitos;**

- Hospedagem vinculada à UFRJ;
- Emissão de certificados.



Gratuito

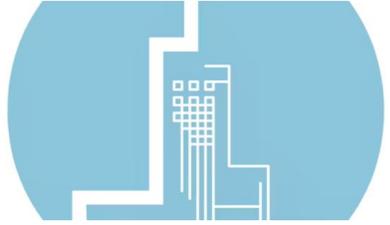


USP Universidade de São Paulo

**Introdução à Ciência da Computação com Python Parte 1**

★ 4,9 · 7,3 mil avaliações

Básico · Curso · 1 a 3 meses



USP Universidade de São Paulo

**Introdução aos Princípios e Práticas da Gestão De Projetos**

★ 4,8 · 1,2 mil avaliações

Básico · Programa de cursos integrados · 3 a 6 meses

Gratuito



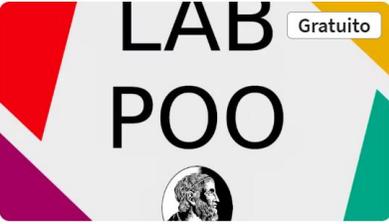
USP Universidade de São Paulo

**Introdução à Ciência da Computação com Python Parte 2**

★ 4,9 · 1,5 mil avaliações

Básico · Curso · 1 a 3 meses

Gratuito



USP Universidade de São Paulo

**Laboratório de Programação Orientada a Objetos - Parte 1**

★ 4,8 · 176 avaliações

Intermediário · Curso · 1 a 3 meses

Gratuito



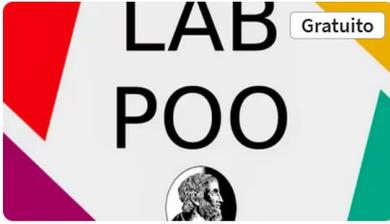
USP Universidade de São Paulo

**Processamento Neural de Linguagem Natural em Português I**

★ 4,8 · 17 avaliações

Intermediário · Curso · 1 a 3 meses

Gratuito

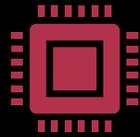


USP Universidade de São Paulo

**Laboratório de Programação Orientada a Objetos - Parte 2**

★ 5 · 17 avaliações

Avançado · Curso · 1 a 3 meses



Introdução à ciência de dados para bioinformática;



Programação em Python para bioinformática

# Universidade

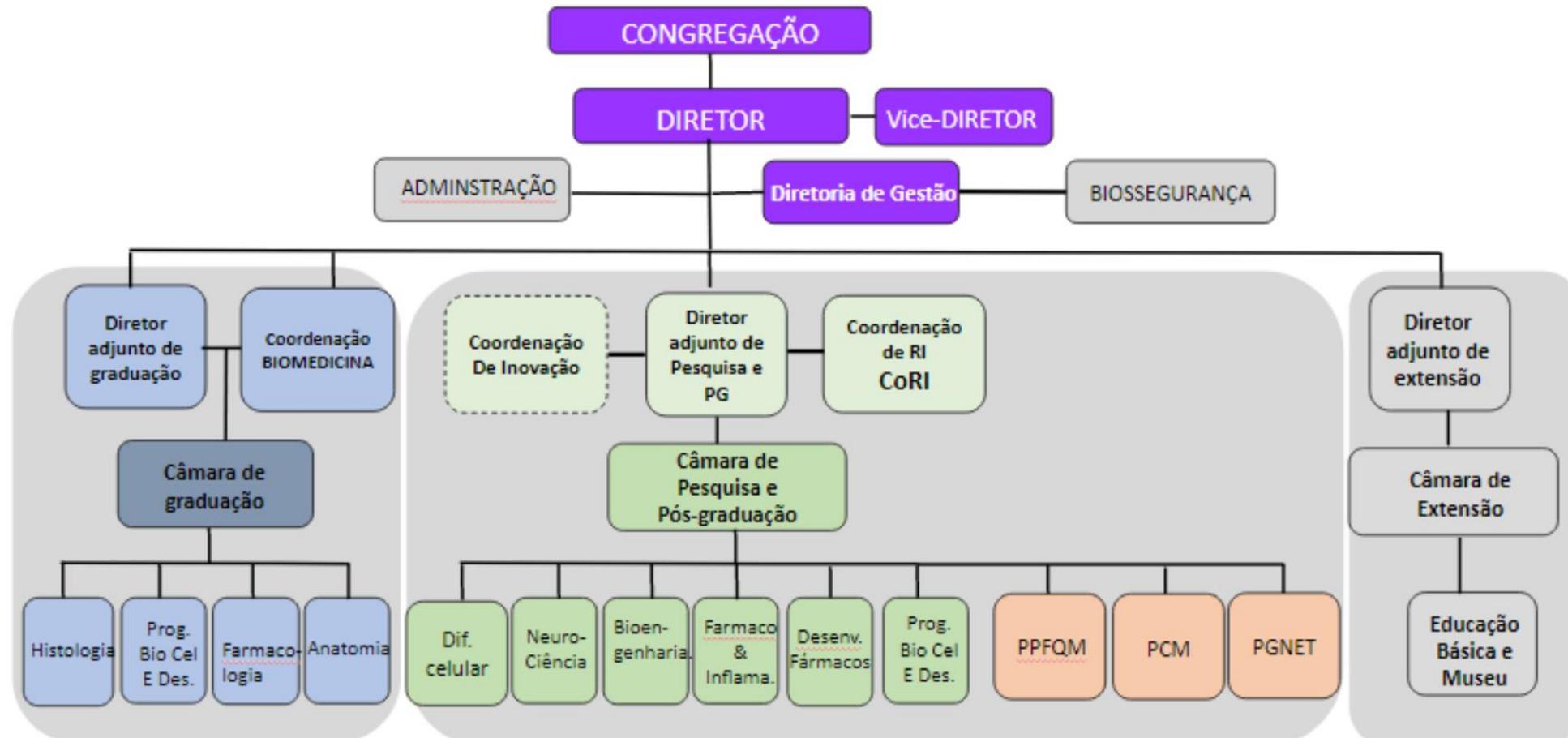
I. Ensino

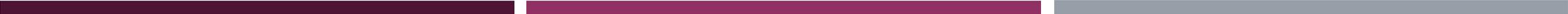
II. Pesquisa

III. Extensão

IV. Administrativa

# ESTRUTURA DO ICB





OBRIGADO!