FUNDAÇÃO GETULIO VARGAS
ESCOLA DE MATEMÁTICA APLICADA – FGV/EMAp
CURSO DE GRADUAÇÃO EM MATEMÁTICA APLICADA

TERMO DE COMPROMISSO PARA ORIENTAÇÃO ACADÊMICA

Nome do Aluno:	Iago Riveiro Santos Dutra
Nome do Professor Orientador:	Alberto Paccanaro
Nome do Professor Tutor:	Paulo Cezar Pinto Carvalho
Título provisório do trabalho:	Comparison of Machine Learning Algorithms for Drug Embedding Generation from SMILES: Evaluating Representation Quality for Downstream Tasks

Proposta do trabalho:

The motivation for this project stems from the wide variety of algorithms available for generating drug embeddings from molecular structures. With so many options, researchers often face difficulty in choosing the most suitable method for their specific needs, leading to inefficient use of time and computational resources.

This study aims to address this challenge by comparing different machine learning algorithms for creating molecular embeddings from SMILES (Simplified Molecular Input Line Entry System) strings. We will evaluate Variational Auto-Encoders (VAEs), Transformers, Graph Neural Networks (GNNs), Message-Passing Graph Neural Networks (MPGNNs), and traditional baseline fingerprints. The quality of the embeddings will be assessed using a dataset of drug side effect frequencies, which will serve as a benchmark to measure how well each model captures relevant molecular features. A fully connected neural network will be used to map the embeddings to final predictions for downstream tasks. Additionally, we will discuss the strengths and limitations of each model, not only for side effect prediction but also for other applications like drug design.

All implementations will be done using the PyTorch library, and the code will be made publicly available on GitHub to ensure reproducibility and encourage further research.

Os abaixo-assinados concordam em desempenhar seus respectivos papéis na orientação do Trabalho de Conclusão acima.

Rio de Janeiro, 10 de março de 2025

Assinatura do Aluno:	
Assinatura do Professor Orientador:	
Assinatura do Professor Tutor:	