



**Pedro Filipe
Carneiro Venâncio**

**Aplicação de novos algoritmos bioinformáticos na
análise de dados de Next Generation Sequencing
(NGS)**

**Application of new bioinformatic algorithms in Next
Generation Sequencing (NGS) data analysis.**

DOCUMENTO PROVISÓRIO



**Pedro Filipe
Carneiro Venâncio**

**Aplicação de novos algoritmos bioinformáticos na
análise de dados de Next Generation Sequencing
(NGS)**

**Application of new bioinformatic algorithms in Next
Generation Sequencing (NGS) data analysis.**

DOCUMENTO PROVISÓRIO

*“The greatest challenge to any thinker is stating the problem in a
way that will allow a solution”*

— Bertrand Russell



**Pedro Filipe
Carneiro Venâncio**

**Aplicação de novos algoritmos bioinformáticos na
análise de dados de Next Generation Sequencing
(NGS)**

**Application of new bioinformatic algorithms in Next
Generation Sequencing (NGS) data analysis.**

Relatório de estágio curricular apresentado à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Bioinformática Clínica, especialização em Bioinformática do Genoma , realizado sob a orientação científica da Doutora Gabriela Maria Ferreira Ribeiro de Moura, Professora auxiliar do Departamento de Ciências Médicas da Universidade de Aveiro, e supervisão da Doutora Alexandra Filipa Lopes, membro da entidade de acolhimento UNILABS.

Dedico este trabalho à minha esposa e filho pelo incansável apoio.

o júri / the jury

presidente / president

Prof. Doutor João Antunes da Silva

professor associado da Faculdade de Engenharia da Universidade do Porto

vogais / examiners committee

Prof. Doutor João Antunes da Silva

professor associado da Faculdade de Engenharia da Universidade do Porto

Prof. Doutor João Antunes da Silva

professor associado da Faculdade de Engenharia da Universidade do Porto

Prof. Doutor João Antunes da Silva

professor associado da Faculdade de Engenharia da Universidade do Porto

Prof. Doutor João Antunes da Silva

professor associado da Faculdade de Engenharia da Universidade do Porto

Prof. Doutor João Antunes da Silva

professor associado da Faculdade de Engenharia da Universidade do Porto

**agradecimentos /
acknowledgements**

Agradeço toda a ajuda a todos os meus colegas e companheiros.

Palavras Chave

NGS, arquitetura, história, construção, materiais de construção, saber tradicional.

Resumo

Um resumo é um pequeno apanhado de um trabalho mais longo (como uma tese, dissertação ou trabalho de pesquisa). O resumo relata de forma concisa os objetivos e resultados da sua pesquisa, para que os leitores saibam exatamente o que se aborda no seu documento.

Embora a estrutura possa variar um pouco dependendo da sua área de estudo, o seu resumo deve descrever o propósito do seu trabalho, os métodos que você usou e as conclusões a que chegou.

Uma maneira comum de estruturar um resumo é usar a estrutura IMRaD. Isso significa:

- Introdução
- Métodos
- Resultados
- Discussão

Veja mais pormenores aqui:

<https://www.scribbr.com/dissertation/abstract/>

Keywords

textbook, architecture, history, construction, construction materials, traditional knowledge.

Abstract

An abstract is a short summary of a longer work (such as a thesis, dissertation or research paper).

The abstract concisely reports the aims and outcomes of your research, so that readers know exactly what your paper is about.

Although the structure may vary slightly depending on your discipline, your abstract should describe the purpose of your work, the methods you've used, and the conclusions you've drawn.

One common way to structure your abstract is to use the IMRaD structure. This stands for:

- Introduction
- Methods
- Results
- Discussion

Check for more details here:

<https://www.scribbr.com/dissertation/abstract/>

Contents

Contents	i
List of Figures	iii
List of Tables	v
Lista de Excertos de Código	vii
Glossário	ix
1 Introduction	1
1.1 Internship Context and Framework	1
1.2 Motivation and Objectives of the Work	2
1.3 Document Structure	2
1.4 Characterization of the Host Entity and Work Plan	2
1.4.1 UNILABS	2
1.4.2 Schedule	2
1.5 Theoretical Framework	2
2 Analysis tool	3
2.1 Internship Context and Framework	3
2.2 Motivation and Objectives of the Work	4
2.3 Document Structure	4
2.4 Characterization of the Host Entity and Work Plan	4
2.4.1 UNILABS	4
2.4.2 Schedule	4
2.5 Theoretical Framework	4
A Additional content	5

List of Figures

List of Tables

Lista de Excertos de Código

Glossário

Introduction

"The only source of knowledge is experience." - Albert Einstein

1.1 INTERNSHIP CONTEXT AND FRAMEWORK

This document represents the final report of the internship carried out as an integral part of the Internship Course (49991) of the second year of studies of the Master's Degree in Clinical Bioinformatics, specializing in Genome Bioinformatics, at the University of Aveiro. The internship lasted for nine months, starting on November 21, 2023, and concluding on July 19, 2024, totaling 1296 hours of work.

During this period, the intern had the opportunity to apply the knowledge acquired throughout the course and to engage in practical projects related to bioinformatics and genomics. UNILABS, a renowned company in the healthcare sector, provided a professional environment where the intern could collaborate with experienced professionals and actively participate in projects relevant to clinical bioinformatics. This report addresses the activities developed during the internship and the contributions to the projects in which the intern was involved.

This introductory section aims to provide an overview of the context in which the internship was carried out, establishing the groundwork for understanding the activities and results presented throughout the report.

- 1.2 MOTIVATION AND OBJECTIVES OF THE WORK
- 1.3 DOCUMENT STRUCTURE
- 1.4 CHARACTERIZATION OF THE HOST ENTITY AND WORK PLAN
 - 1.4.1 UNILABS**
 - 1.4.2 Schedule**
- 1.5 THEORETICAL FRAMEWORK

Analysis tool

"The only source of knowledge is experience." - Albert Einstein

2.1 INTERNSHIP CONTEXT AND FRAMEWORK

This document represents the final report of the internship carried out as an integral part of the Internship Course (49991) of the second year of studies of the Master's Degree in Clinical Bioinformatics, specializing in Genome Bioinformatics, at the University of Aveiro. The internship lasted for nine months, starting on November 21, 2023, and concluding on July 19, 2024, totaling 1296 hours of work.

During this period, the intern had the opportunity to apply the knowledge acquired throughout the course and to engage in practical projects related to bioinformatics and genomics. UNILABS, a renowned company in the healthcare sector, provided a professional environment where the intern could collaborate with experienced professionals and actively participate in projects relevant to clinical bioinformatics. This report addresses the activities developed during the internship and the contributions to the projects in which the intern was involved.

This introductory section aims to provide an overview of the context in which the internship was carried out, establishing the groundwork for understanding the activities and results presented throughout the report.

2.2 MOTIVATION AND OBJECTIVES OF THE WORK

2.3 DOCUMENT STRUCTURE

2.4 CHARACTERIZATION OF THE HOST ENTITY AND WORK PLAN

2.4.1 UNILABS

2.4.2 Schedule

2.5 THEORETICAL FRAMEWORK

APPENDIX A

Additional content