

Relatório Final de Estágio Supervisionado

por

Mateus de Araujo Silva

Orientador: Profa. Pollyana de Queiroz Ribeiro

Santa Helena de Goiás, 31 de março de 2022

RELATÓRIO DE ESTÁGIO SUPERVISIONADO

Nome do Estagiário: Mateus de Araujo Silva

Matrícula: 12019006023

Período que estuda: 6º período

Nome da Empresa: Unidade Universitária de Santa Helena de Goiás

Nome do Supervisor Local: Adriano Ferraz da Costa

Área de atuação: Ciência de dados orientados a grafos

Período de duração: Data de Início: 25/01/2022

Data de Conclusão: 18/03/2022

Horas Semanais: 15 Total de Horas: 90

31/03/2022
Data da Entrega

Assinatura do Aluno

Visto do Professor

SUMÁRIO

1 INTRODUÇÃO	4
2 APRESENTAÇÃO DA EMPRESA	4
3 DESCRIÇÃO DO ESTÁGIO	5
4 CONCLUSÃO	8
ANEXOS	9

1 INTRODUÇÃO

O Estágio Supervisionado submete-se às determinações contidas nas seguintes legislações:

Lei Federal nº 11.788/2008.

Diretrizes Curriculares Nacionais dos Cursos - DCNs.

Regimento Geral da Universidade Estadual de Goiás.

A Universidade Estadual de Goiás é uma instituição de ensino, pesquisa e extensão, de caráter público, gratuito e laico, composta por estrutura multicampi, com âmbito de atuação em todo o território do Estado de Goiás, dotada de autonomia didático-científica, administrativa, de gestão financeira e patrimonial.

Este relatório aborda o desenvolvimento da disciplina Estágio Supervisionado IV, a apresentação da organização concedente e as atividades realizadas.

2 APRESENTAÇÃO DA EMPRESA

A UEG tem como missão produzir e socializar o conhecimento científico e o saber, desenvolver a cultura e formação integral de profissionais e indivíduos capazes de se inserirem criticamente na sociedade e assim promoverem a transformação da realidade socioeconômica do Estado de Goiás e do Brasil. Na condição de instituição multicampi, a UEG possui Câmpus Universitários em todas as microrregiões do Estado.

O Estágio Supervisionado aconteceu sob a supervisão do Prof. Adriano Ferraz, na Unidade Universitária de Santa Helena de Goiás, pertencente ao Campus Regional Sudoeste com sede na cidade de Quirinópolis.

3 DESCRIÇÃO DO ESTÁGIO

As atividades desenvolvidas estão no segmento de Ciência de dados do ramo de atividades de pesquisa, que fazem parte dos pontos de interesse do alinhamento estratégico da organização em incentivar a inovação científica e tecnológica, o empreendedorismo e a transferência de tecnologia, o desenvolvimento e difusão do aprimoramento acadêmico e cultural, e o fortalecimento da interação da UEG com a sociedade.

Na execução das atividades foram utilizadas abordagens multidisciplinares como leituras sobre conceitos de teoria de grafos com o objetivo de revisar parte do conteúdo estudado anteriormente na disciplina de teoria de grafos, como problemas notáveis onde soluções envolvendo teoria dos grafos são aplicadas, as definições dos termos mais utilizados em referência a grafos e os tipos de grafos - esses conteúdos foram alcançados por interesse próprio, principalmente através do artigo da wikipédia *Graph Theory* e seus artigos correlatos. Encerrada a revisão, um dos artigos correlatos é *Network Theory* que trata de análise de rede e aborda métricas para descrição das redes, otimização e resoluções de problemas, considerando principalmente o relacionamento entre os nós. Parte do mesmo conteúdo também é abordado no conteúdo indicado pelo supervisor, que consiste do material fornecido pelo site Neo4j *Graph Academy* e biblioteca algoritmos de Ciência de dados orientados a grafos da plataforma Neo4j. O material indicado pelo supervisor trata o assunto de análise de rede mais profundamente abordando matemática e principalmente aplicação prática, que foi o foco das atividades de estágio que são abordadas a seguir.

Os conjuntos de dados utilizados foram *Game of Thrones* (multigrafo composto por nós que representam os personagens do seriado de televisão e relacionamentos com peso que representam a interação entre os personagens nas temporadas 1 a 7, contando com 400 nós e 7102 relacionamentos), *European Roads* (grafo composto por nós que representam cidades do continente europeu e relacionamentos que representam estradas entre as cidades, contando com 894 nós e 2499 relacionamentos), e *Recipes* (grafo composto por nós que representam receitas, nós que representam ingredientes e relacionamentos que indicam quais são os ingredientes utilizados na receita ou ingrediente, contando com 3562 nós e 20476 relacionamentos), todos conjuntos de dados são fornecidos pela plataforma Neo4j e foram escolhidos por serem os conjuntos de dados utilizados no guia de introdução à algoritmos de grafos.

Foram executadas atividades práticas de implementação de algoritmos de Ciência de dados em bancos de dados orientados a grafos com o uso da linguagem *Cypher* na plataforma Neo4j *Desktop*, incluindo estudos sobre a teoria e propósito, que são abordados no guia de curso de introdução a algoritmos de grafos do site Neo4j *Graph Academy* e sintaxe dos algoritmos, que é abordada na biblioteca de algoritmos Ciência de dados orientados a grafos da plataforma Neo4j. Os algoritmos utilizados foram: Componentes fracamente conectados (algoritmo para detectar subgrafos desconexos na rede, aplicado no conjunto de dados *European Roads* com o objetivo de detectar conjuntos de cidades isoladas levando em consideração a inexistência de estradas que as conecte a outros conjuntos de cidades); Propagação de rótulo (algoritmo para descrever comunidades se baseando na quantidade de nós vizinhos, aplicado no conjunto de dados *European Roads* com o objetivo de descrever comunidades levando em consideração a densidade de conexões que as cidades têm entre si, que são representações das estradas); Modularidade de *Louvain* (algoritmo para detectar comunidades baseado no método de *louvain*, aplicado nos conjuntos de dados *European Roads* e *Recipes* com o objetivo de detectar comunidades baseado no método de *Louvain*); Contagem de triângulos (algoritmo para detectar triângulos, que são subgrafos compostos por três nós onde cada nó está conectado aos outros dois, aplicado no conjunto de dados *European Roads* para descobrir quais cidades estão envolvidas em mais triângulos); Coeficiente de agrupamento local (algoritmo para estimar a probabilidade de que nós vizinhos a um dado nó estejam relacionados entre si, aplicado no conjunto de dados *European Roads*); *PageRank* (algoritmo para estimar a influência de cada nó levando em consideração quantos nós se referem a ele e a influência de seus nós vizinhos, aplicado no conjunto de dados *European Roads* para descobrir a influência das cidades considerando suas estradas e as estradas das cidades vizinhas, e aplicado no conjunto de dados *Recipes* para identificar os ingredientes mais influentes); Centralidade de intermediação (algoritmo para detectar a influência de um nó para o fluxo da rede, aplicado no conjunto de dados *European Roads* para detectar cidades por onde pessoas viajando de carro provavelmente vão passar, considerando a frequência dessas cidades em caminhos dado qualquer par de cidades); Similaridade de nós (algoritmo para denotar o quanto dados nós são parecidos levando em consideração os nós a que eles estão relacionados, aplicado no conjunto de dados *Recipes* com o objetivo de identificar receitas que utilizam os mesmos ingredientes);

Também foram executadas atividades práticas com a interface de visualização de dados NEuler que permite executar os mesmos algoritmos já descritos, porém utilizado

interface gráfica e botões ao invés de escrever consultas em linguagem *Cypher*. Nestas atividades foi utilizado o conjunto de dados *Game of Thrones* pela familiaridade do estagiário com o conhecimento que é representado pelo grafo; E atividades de projeção de banco de dados, que abordam como é carregado o banco de dados ou parte selecionada do banco de dados que está na memória secundária para a memória principal;

A execução das atividades foi interrompida temporariamente devido a motivos de saúde do estagiário, justificado por meio de atestado médico (5 dias). A previsão para a conclusão das atividades e entrega do último relatório parcial foi dia 08/03/2022 e devido ao motivo citado as atividades foram efetivamente concluídas dia 18/03/2022.

4 CONCLUSÃO

O Estágio Supervisionado é uma importante parte integradora do currículo dos cursos de graduação. Por meio deste componente curricular, o acadêmico entra em contato direto com sua realidade profissional, tomando-a como objeto de estudo, de análise e investigação, a fim de concretizar e aprimorar os conhecimentos adquiridos e consolidar o processo de profissionalização.

Agradeço a UEG e Unidade Universitária de Santa Helena de Goiás por me conceder a oportunidade de realização dessa atividade e aos professores Adriano Ferraz e Pollyana Queiroz pelo apoio e instrução.

ANEXOS

Termo de compromisso e Relatório Final (Estágio Supervisionado IV)

<https://github.com/SilvaMateusdeAraujo/estagio-supervisionado-ueg/>

Neo4j Graph Academy

<https://neo4j.com/graphacademy/>

Biblioteca algoritmos de Ciência de dados orientados a grafos Neo4j

<https://neo4j.com/developer/graph-data-science/graph-algorithms/>

Versão	Data de entrega
Relatório final	18/03/2022
Revisão	26/03/2022
Relatório final com adequações	28/03/2022
Revisão 2	30/03/2022
Relatório final com adequações 2	31/03/2022