



NOVO MODELO DE RELATÓRIO DOS BOLSISTAS

09/2021

Dados da Bolsa

Tipo de Bolsa: ☒ IC ☐ TCC ☐ PqEP ☐ ME

Nome do/a Orientador/a: Anna Helena Reali Costa

Nome do Projeto: Democracia Aumentada

Período da Bolsa: 02/02/2022 a 01/02/2023

Relatório: ☐ Final ☒ Parcial

Período do Relatório: 25/03/2022 a 25/04/2022

Descrição das Atividades de Pesquisa do Projeto

Descrição das atividades acadêmicas: 2º Semestre Letivo, em que o aluno está cursando as seguintes matérias:

- PEA3306 - Conversão Eletromecânica de Energia
- PEA3301 - Introdução aos Sistemas de Potência
- PEA3311 - Laboratório de Conversão Eletromecânica de Energia
- PTC3307 - Sistemas e Sinais
- PCS3335 - Laboratório Digital A
- PSI3213 - Circuitos Elétricos II
- PEA3100 - Energia, Meio Ambiente e Sustentabilidade
- PMT3100 - Fundamentos de Ciência e Engenharia dos Materiais
- PMT3131 - Química dos Materiais Aplicada à Engenharia Elétrica

Descrição das atividades planejadas para o relatório (repetir do relatório anterior):

Para o próximo mês é planejado estudar, implementar e iniciar os testes do ZeroBERTo em outros benchmarks.

Descrição das atividades de pesquisa realizadas:

Durante o mês foram discutidos os próximos passos para o desenvolvimento do ZeroBERTo em questão de arquitetura e implementação, quanto a isso vamos utilizar técnicas de *entailment* e *contradiction* para ter um



modelo melhor e de uso mais geral, expandiremos para multilingual e também para a tarefa de classificação multiclasse. Esses foram os objetivos definidos durante as reuniões.

Além disso, fizemos uma grande revisão de artigos sobre *Contrastive Learning* e decidimos implementar uma versão em português do artigo “[SimCSE: Simple Contrastive Learning of Sentence Embeddings](#)”, que foi pautado como um dos estado-da-arte em *embeddings*. A ideia é criar um dos melhores *embeddings* para o português usando o BERTimbau como base, aliado ao método de *contrastive learning* (não-supervisionado) para criação de *embeddings* de frases que possam ser usados por outros pesquisadores em *machine learning*. Esses *embeddings* servirão de base para criação do ZeroBERTo e são planejadas 2 escritas de artigos, uma para cada um desses objetivos.

No que diz respeito à criação da interface do Democracia Aumentada é planejado continuar com os dados do DAR, porém por conta da alteração no cronograma o bolsista está estudando a implementação do site interativo através da própria plataforma do *HuggingFace Transformers*, em que se pode usar os modelos que já existem no site, bem como escrever códigos de forma mais direta. Essa proposta substitui o plano inicial de se trabalhar com Django e um desenvolvimento web “do zero”, o que demandaria o tempo previsto originalmente na bolsa.

Descrição das próximas atividades:

Para o próximo mês é planejado já ter o código funcional para geração dos *embeddings* que será, em sua versão final, executado em um dos computadores do C2D usando o corpus Carolina.

Houve alteração significativa no tema ou prazo: (X) Sim () Não

Justifique em caso positivo:

A mudança se deve por conta que o bolsista foi selecionado para um processo de intercâmbio para Alemanha. Como as aulas na RTWH Aachen University serão iniciadas no início de outubro, um novo cronograma foi feito, visando a finalização da bolsa atual antes do final do período inicialmente estipulado.

Mês	Atividade
Maio	Implementação do <i>sentence embedding</i> e aplicação no ZeroBERTo
Junho	Escrita de dois artigos sobre os feitos de maio
Julho	Rotulagem de todos os dados do DAR usando o ZeroBERTo e criação dos filtros
Agosto	Implementação da interface e outros módulos interativos na plataforma do <i>HuggingFace</i>



UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO
DEPARTAMENTO DE ENGENHARIA DE COMPUTAÇÃO E SISTEMAS DIGITAIS
CENTRO DE CIÊNCIA DE DADOS (C²D)



Setembro	Apresentação do trabalho no SIICUSP e escrita do relatório final do C2D e da IC no sistema Atena da USP
Apreciação Circunstanciada do/a Orientador/a sobre as Atividades da/o Bolsista	
Etapas cumpridas no relatório: () Ótimo (X) Bom () Regular () Fraco	
Programação para a próxima etapa: (X) Ótimo () Bom () Regular () Fraco	
Resultados em relação às expectativas iniciais: () Acima (X) Dentro () Abaixo () Muito abaixo	
Previsão de conclusão no prazo: () Sim (X) Não	
Justifique em caso negativo: Vide justificativas e novo cronograma apresentados acima.	
Apreciação da orientadora: O aluno tem sido muito participativo e com boas contribuições. Devido à sua seleção para realização de um intercâmbio na Alemanha, alteramos seu cronograma para que finalize a IC com sucesso e gere publicações resultantes das pesquisas.	
Protocolo	
Data: 25/04/2022	
Nome Completo da/o Bolsista: Enzo Bustos Da Silva	