

Dados da Bolsa
Tipo de Bolsa: <input checked="" type="checkbox"/> IC <input type="checkbox"/> PqEP <input type="checkbox"/> TCC <input type="checkbox"/> ME <input type="checkbox"/> DO <input type="checkbox"/> PD
Nome do Orientador/Supervisor: Anna Helena Reali Costa
Nome do Projeto: Chatbot Q&A multi-agente
Período da Bolsa: 01/02/2021 a 01/02/2022
Relatório: <input type="checkbox"/> Final <input checked="" type="checkbox"/> Parcial
Período do Relatório: 25/01/2021 a 25/02/2021
Descrição das Atividades de Pesquisa do Projeto
Descrição das atividades acadêmicas: Período de Férias Letivas
Descrição das atividades de pesquisa: Durante este período foram iniciadas as atividades de pesquisa na área de NLP e NLU, voltadas principalmente para a revisão bibliográfica e leitura de artigos sobre os princípios e métodos usados nesta área como embeddings (Doc2Vec e GloVe), tokenizers, lemmatization e stemming. Buscou-se por trabalhos, códigos e bibliotecas que já implementam essas técnicas, a exemplo de bibliotecas temos o SpaCy, Stanza, NLTK, ftfy e gensim. Também foi iniciado um estudo em expressões regulares (Regex) que é útil para trabalhar com textos. Foi utilizado um dataset conhecido de avaliação de filmes, o iMDB dataset para testar em pequena escala tanto alguns desses conceitos como também as bibliotecas que foram citadas. Além disso, foram também pesquisados métodos de extração de textos de PDFs para criar um corpus de dados, testando ferramentas como pdfminer.six, Py2PDF e pdfplumber. Finalmente, foram realizadas reuniões semanais com o restante do grupo atuante em NLP, para acompanhamento das atividades.
Houve alteração no cronograma original: <input type="checkbox"/> Sim <input checked="" type="checkbox"/> Não
Justifique em caso positivo:
Apreciação Circunstanciada do Orientador/Supervisor sobre as Atividades do Bolsista
Apreciação:
Etapa cumprida no relatório: <input type="checkbox"/> Ótimo <input checked="" type="checkbox"/> Bom <input type="checkbox"/> Regular <input type="checkbox"/> Fraco
Programação para a próxima etapa: <input type="checkbox"/> Ótimo <input checked="" type="checkbox"/> Bom <input type="checkbox"/> Regular <input type="checkbox"/> Fraco
Resultados em relação às expectativas iniciais: <input type="checkbox"/> Acima <input checked="" type="checkbox"/> Dentro <input type="checkbox"/> Aquém <input type="checkbox"/> Muito aquém
Previsão de conclusão no prazo: <input checked="" type="checkbox"/> Sim <input type="checkbox"/> Não
Justifique em caso negativo:
Comentários: O aluno iniciou seus estudos e está tomando conhecimento dos conceitos fundamentais e ferramentas usualmente utilizadas no processamento de textos. Tem participado de reuniões e mostrado um perfil pró-ativo.
Protocolo
Data: 25/02/2021
Nome Completo do Bolsista: Enzo Bustos Da Silva