

Turing: Chatbot para auxílio e detecção de ineteressas na área de computação

Title: Turing: Chatbot for assistance and detection of interest in the computing area

<Gabriel Di iorio Silva>

<Universidade Federal de Juiz de Fora (UFJF)>

<iorio@ice.ufjf.br>

<Gabriel Albuquerque de Souza Hosken Magalhães>

<Universidade Federal de Juiz de Fora (UFJF)>

<gabriel.albuquerque@ice.ufjf.br>

<Júlia de Almeida Valadares>

<Universidade Federal de Juiz de Fora (UFJF)>

<juliavaladares@ice.ufjf.br>

Resumo

<Here comes the abstract of the paper in Portuguese or Spanish, if that's the language of the manuscript. The abstract should summarize the contents of the manuscript and should contain at least 150 and at most 300 words long and must be written in italics, Times 10, justified, with no special indentation and no spacing before or after.>

Palavras-chave: *Abstract must be followed by 3 to 10 keywords. The keywords should be justified with a line space single, no special indentation, with no spacing before and spacing of exactly 24-points after. The text should be set in Times 10-point font size and in italic font style. Please use semi-colon as a separator. Keywords must be title cased.>*

Abstract

<Here comes the abstract of the paper (in English). The abstract should summarize the contents of the manuscript and should contain at least 150 and at most 300 words long and must be written in italics, Times 10, justified, with no special indentation and no spacing before or after.>

Keywords: *Abstract must be followed by 3 to 10 keywords. The keywords should be justified with a line space single, no special indentation, with no spacing before and spacing of exactly 24-points after. The text should be set in Times 10-point font size and in italic font style. Please use semi-colon as a separator. Keywords must be title cased.>*

1 Introdução

Com a instauração da internet como uma ferramenta de trabalho geral em todos os âmbitos do cotidiano do cidadão moderno, diversas perspectivas de trabalho, negócios, educação e divulgação, por exemplo, ganharam novas formas. De forma simultânea com este avanço, as reduções dos tempos de resposta, a conectividade até nos lugares mais remotos e a redução de tamanho físico dos componentes mecânicos dos computadores auxiliaram para que esse fenômeno ganhasse dimensões jamais antes vistas para a população. Com estes alicerces, o termo de Transformação Digital ganhou destaque como um termo para relatar a adoção de tecnologias disruptivas para aumentar a produtividade, valorização e bem-estar social (Ebert & Duarte, 2018).

Para entendermos melhor compreensão de tal aspecto é interessante uma forma de visualização que apresente o impacto da Transformação Digital que nas abrangentes esferas sociais e áreas. Logo, podemos destacar o Vortex de Transformação Digital que nos apresentam em um estudo, anual, em que colocação a área apresentada ocupa dentre aquelas que sofrem impacto do fenômeno supracitado. Seguindo o conceito de Vortex, aquelas que sofrem mais próximas do seu centro podem ser designadas como as que sofrem maior impacto com a Transformação digital e mais externas as que sofrem menos, mas que ainda sim carregam bastante influência:

Na figura 1 é possível perceber que, dentre as áreas, a educação ganha um espaço de destaque ocupando a 8 colocação no estudo de 2019. Com esse fato podemos compreender como a área do ensino vem sofrendo mudanças significativas no âmbito de implementar cada vez mais tecnologias nas suas formas e técnicas de aprendizagem. Uma dessas formas que toma forma e mais destaque atualmente se refere à Gamificação. Este termo pode ser compreendido como as práticas destinadas ao processo de aprendizagem que buscam estimular o engajamento do indivíduo a fim de proporcionar experiências mais efetivas e relevantes ao sujeito por meio de jogos e técnicas de recompensa aplicadas e jogo (Busarello, 2016).

Logo, inseridos no contexto de Transformação Digital e munidos com o conhecimento da Gamificação a questão de como criar sistemas e jogos com os conceitos deste último que cativem, intriguem e despertem o interesse em busca do aprendizado são de grande valia para auxiliar na composição das técnicas de aprendizagem.

1.1 Uso da Gamificação ou Jogos no Ensino e Aprendizagem de Computação

Diversas são as mecânicas para o uso da gamificação no processo de ensino, alguns delas são destacadas e apontadas no trabalho de (Nah, Zeng, Telaprolu, Ayyappa, & Eschenbrenner, 2014). Nele podemos perceber ferramentais com o objetivo motivacional, de engajamento, status e performance. Dessa forma podemos ter dimensão de como abordar essa temática onde todas visam o mesmo fim pois, despertar e tornar essa atividade educativa e gamificada simultaneamente exige o emprego de técnicas diferentes para auxiliar na compreensão daquelas mais efetivas e como foram aplicadas. Para tal, devido à sua enorme extensão, trabalhos como (Huang & Soman, 2013) relatam em forma de guia alguns passos para abstração e visualização da problemática a ser

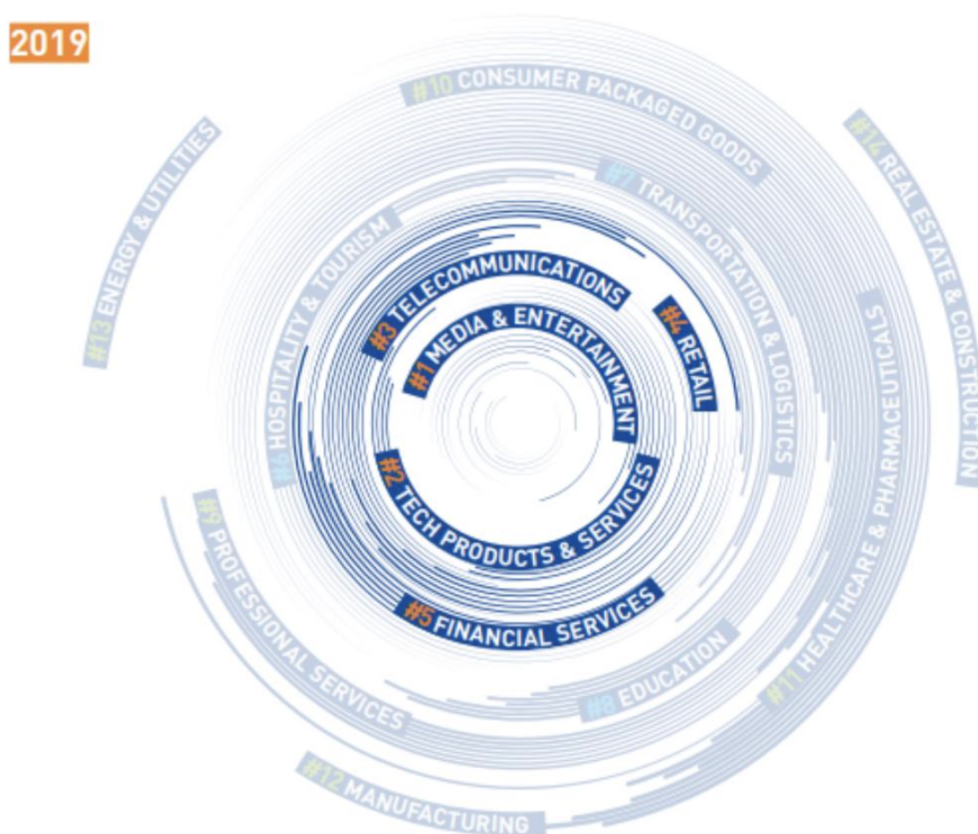


Figura 1: Vortex de Transformação Digital.

tratada. Fatores importantes que são apontado são: Entender o público alvo e o contexto, Definir os objetivos de aprendizagem, Estruturar a experiência, Identificar os recursos e Aplicar os elementos de gamificação. Cada etapa deve ser atentamente trabalhada pela equipe para que a sua proposta consiga ensinar o que almeja com a melhor experiência possível que cativa o público alvo. Logo, muito além do que simplesmente construir um sistema ou jogo para o ensino, a gamificação aplicada à educação pode ser compreendida como uma maneira de entender as interações sociais e métodos de aprendizagem concomitantemente. Todas as informações aqui ressaltadas podem ser aplicadas, também, para o conceito de gamificação associado ao ensino de aspectos relacionados aos cursos de computação. Portanto, nesse trabalho apresentamos a proposta **Turing**. Um chatbot que tem como objetivo a elucidação dos alunos na busca por encontrar sua área de interesse dentre as diversas presentes na computação como também auxiliá-lo com materiais didáticos e dúvidas sobre cursos e suas diferenças que, por diversas vezes, são recorrentes. Em vista disso, apresenta-se como diferencial desse trabalho o uso de técnicas de gamificação ligadas à conceitos de sistemas de recomendação para que o trajeto no intuito de compreender melhor as áreas da computação e quais são mais inerentes à personalidade do aluno avaliado.

O trabalho encontra-se subdividido como segue: Na seção 1 é apresentada a introdução.....

2 Trabalhos Relacionados

Na busca por trabalhos relacionados e para compreensão da gamificação aplicada a alunos de computação. Primeiramente foi necessário entender os benefícios desse fator dentro da metodologia de aprendizagem e resultados comparativos para que assim a proposta pudesse ser elaborada de forma eficiente e com a dimensão de toda a ampliação de conhecimento por ela gerada. Assim, alguns trabalhos exibem resultados de como o engajamento dos estudantes teve uma melhora significativa em estudantes que foram apresentados à técnicas de gamificação quando comparados com os que não foram. Entretanto, nenhuma melhora significativa foi notada no âmbito de performance (Ortiz Rojas, Chiliza, & Valcke, 2017). Por conseguinte, formas de aumentar o engajamento dos estudantes foi um dos pilares para este projeto de pesquisa com o intuito de expandir o envolvimento do aluno na busca por sua área de interesse para que a motivação de estudar surja naturalmente.

Além disso, para um amplo entendimento sobre como os sistemas de recomendação designados à suporte estudantil são aplicados, trabalhos que apresentam um panorama dessa área foram amplamente usados afim de observar tendências e formar eficazes de aplicar esse ferramental na educação (Drachsler, Verbert, Santos, & Manouselis, 2015). Neste trabalho os sistemas de recomendações aplicados à educação são agrupados em 7 tipos diferentes: Recursos de recomendação para aprendizado baseados em filtragem colaborativa, Melhorando os algoritmos de filtragem colaborativa com aprendizagem aprimorada, Explorando técnicas de filtragem não colaborativa para recomendações educacionais com êxito, Restrições educacionais como fonte de informação para o processo de recomendação, Informações de contexto para o processo de recomendação, Avaliar o impacto educacional das recomendações em cenários educacionais e Recomendando cursos. Este último agrupamento, em particular, nos auxiliou à obter uma visão geral do projeto e expandi-lo para mais possíveis recomendações além de cursos.

Ademais, trabalhos como o encontrado em (Hiremath, Hajare, Bhosale, Nanaware, & Wagh, 2018), (Clarizia, Colace, Lombardi, Pascale, & Santaniello, 2018) e foram de grande valia para a percepção de formas de apresentar um chatbot para o ambiente educacional e possíveis conceitos a serem aplicados, uma vez que seu trabalho de pesquisa caminha em consonância com os princípios do projeto aqui proposto. Tais projetos debatem sobre as diferentes formas de aplicar um chatbot no âmbito educacional, bem como respostas automáticas a pergunta dos alunos sobre os serviços prestados em nome da equipe acadêmica.

Outrossim, vale destacar o trabalho realizado em (Akcora et al., 2018) é apresentado também um chatbot onde o objetivo não é ter este como um especialista em determinado assunto, mas sim em um corpo de conteúdo que abrange aquele aspecto para que dessa forma seja possível sugerir caminhos ideais entre os itens de conteúdo. Tal trabalho é interessante uma vez que apresenta uma outra forma de abordagem para chatbots em domínio educacional para, ao invés de dar enfoque para um assunto, este é tratado a fim de se agregar assuntos diversificados mas de inúmeros tópicos.

Outra abordagem que serviu para auxiliar na dimensão das possibilidades de ajuda que um chatbot oferece foi encontrada na pesquisa realizada por (Kowsher et al., 2019) em que tal tecnologia é aplicada para educação na língua Bengali que utilizou como método de comparação de frases a técnica de similaridade de Jaccard e nos direcionou em possíveis técnicas para comparação das frases digitadas pelo usuário em relação as encontradas na biblioteca e treinamento do chatbot.

Aliás, para compreender o funcionamento e verificar técnicas de recomendação de materiais para estudantes, projetos de pesquisa como (Salehi, Kamalabadi, & Ghouschi, 2014) tiveram grande serventia uma vez que este apresenta técnicas de recomendação baseado em filtragem colaborativa. Entretanto, apesar de não ser efetivamente uma tecnologia de chatbot, obtivemos um panorama dos possíveis materiais a serem recomendados para estudantes.

Estes projetos foram de suma importância no desenvolvimento do projeto de pesquisa aqui realizado, uma vez que auxiliaram a verificar novas propostas pouco exploradas ao passo que também poderíamos usar trabalhos já consolidados para desenvolver uma proposta inicial servindo como embasamento para essa elaboração. Logo, temos o rudimento bem definido e estudado para todo o processo de desenvolvimento de ideias sobre áreas a se perscrutar dentro das concepções avaliadas

3 Turing: Modelo de chatbot para auxiliar no reconhecimento de interesses em computação

Para realizar, com o auxílio da gamificação, a tarefa de empenhar o aluno na busca por suas áreas de interesse dentro dos diferentes ramos do conhecimento presentes em computação, o modelo de trabalho aqui proposto apresenta um chatbot que visa compreender gostos do aluno em uma conversa rápida e assim reconhecer afinidades para a recomendação de materiais que possam ser úteis ao aluno no seu processo de escolha e descoberta. Para melhor visualização da visão geral do projeto bem como a intelecção de cada uma das particularidades de cada etapa, podemos subdividir este em camadas a fim de descrevê-las de forma detalhada. Para tal, as camadas presentes no sistema Turing são: **Camada de interface** e **Camada de auxílio**.

3.1 Camada de interface

Nesta camada estão compreendidos todos os funcionamentos presentes no modelo de projeto que tem por função realizar todos os estágios que precedem o diálogo com o chatbot e, também, viabilizam a conversa. Para o uso do software é necessário que o usuário realize seu cadastro na plataforma com algumas informações para que seu registro seja salvo dentro do sistema. Com suas informações e registro feito o usuário do tipo aluno já está apto para usufruir do software e é capaz de: Iniciar uma conversa com o chatbot e atualizar suas informações no sistema.

Além disso, é possível também realizar um cadastro de usuário tipo administrador. Com

este perfil o utilizador é capaz de: Editar suas informações de cadastro e cadastrar materiais que serão usados na recomendação. Os materiais cadastrados estão presentes em uma base de dados feita com Postgresql. A partir da percepção da área de interesse do aluno, o chatbot é capaz de recomendar discentes da UFJF especialistas na área, sites com informações sobre o ramo notado, links para vídeos que comentam sobre o campo, livros que também contenham informações sobre o domínio desejado e matérias presentes na grade principal do curso ou em matérias optativas e ou eletivas do departamento de ciência da computação.

Para realizar tal comunicação entre a base de dados e o chatbot uma API (application programming interface) foi desenvolvida a fim de retornar o material cadastrado para o usuário. Essa aplicação reconhece o domínio de interesse por meio de uma tag responsável por caracterizar e assim enviar do material a partir dessa tag. Logo, somente materiais semelhantes ao gosto daquele usuário serão enviados para o aluno. Toda a etapa de comunicação por meio da API é desenvolvida com Psycopg2, um adaptador de Postgresql para a linguagem de programação Python. A escolha deste se deve ao fato de ser o mais utilizado pela comunidade no desenvolvimento de projetos que usufruem desses formatos de base de dados e linguagem de programação. Logo, um maior suporte para mudanças, problemas e otimização do software uma vez que todo o projeto será *open source* para o público que se interessar.

Ademais, no desenvolvimento da interface gráfica do sistema que irá reger o chatbot, foi feito um trabalho de identidade visual montando uma paleta de cores baseada na logo da UFJF com a extensão de navegador ColorZilla a fim de simular um sistema oficial usado na Universidade Federal de Juiz de Fora. A elaboração desta foi feita com o intuito de ser minimalista e intuitiva para os alunos durante seu uso de forma a facilitar o acesso ao chatbot.

3.2 Camada de auxílio

Nesta camada está compreendido todas as funcionalidades relacionadas ao chatbot, sua forma de treinamento, conversação, interpretação e respostas no período de diálogo com este. Com o cadastro realizado e após passar por toda a interface do sistema o mesmo pode iniciar uma comunicação com Turing.

Durante o diálogo, por meio de uma biblioteca de possíveis frases que o usuário pode digitar para se comunicar, o chatbot consegue reconhecer tal frase e assim responder de forma adequada com as respostas também cadastradas na biblioteca. Ao ser iniciado o chatbot é treinado com as bibliotecas que foram criadas para ele e assim sabe como se comunicar diante da situação. Entretanto, como existem diferentes tipos de diálogo, como: Diálogos introdutórios, diálogos a respeito da área de interesse e diálogos de conclusão. Logo, para interpretação e fluidez da comunicação devem ser cadastradas diferentes bibliotecas de diálogo. Dessa forma, foram cadastradas diferentes bibliotecas para interpretação:

- **Biblioteca de cumprimentos:** Está relacionada à possíveis frases que o usuário pode dizer com o intuito de iniciar uma conversa seja essa com palavras formais, linguagem mais

descontraída, perguntas e outras formas de início de diálogo.

- **Bibliotecas de áreas de interesse:** Está relacionada à possíveis frases que indicam gostos que direcionam o chatbot à responder com materiais correspondentes ao interesse. Para cada área de interesse que o sistema será capaz de reconhecer foi cadastrada uma biblioteca a parte.
- **Biblioteca de conclusão:** Está relacionada à possíveis frases que indicam a finalização de uma conversa entre o usuário e o chatbot e denotam que este conseguiu as recomendações necessárias.

Dessa forma, com as bibliotecas implementadas para treinamento do chatbot, todo o funcionamento entre o reconhecimento da frase até a resposta com a sentença correspondente e adequada passa por alguns algoritmos de percepção e de escolhas. Assim, para melhor compreensão e visualização do funcionamento geral do chatbot. Podemos ilustrar seu fluxo de comportamento conforme a figura 2

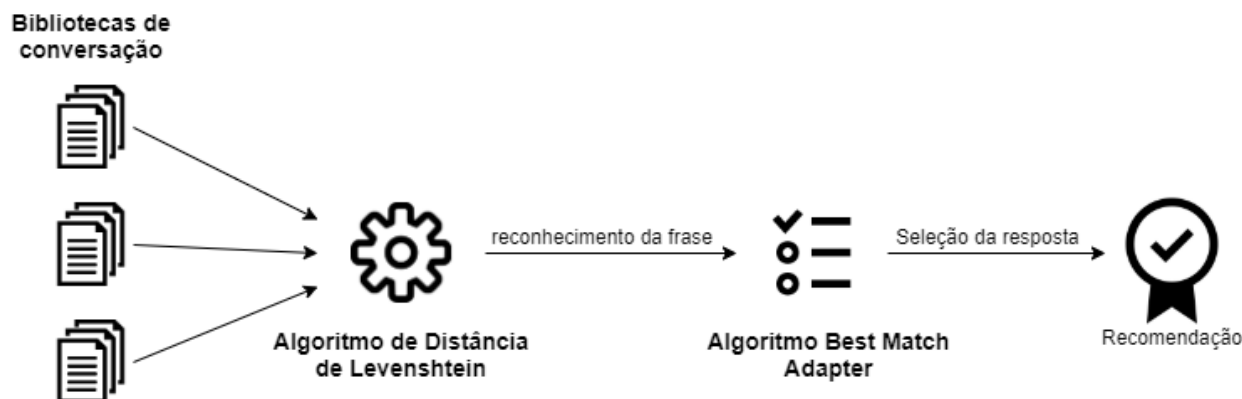


Figura 2: Visão Geral.

Após o treinamento do chatbot com as bibliotecas de conversação cadastradas a frase é encaminhada para o algoritmo de distância de Levenshtein. Nele o chatbot detecta e realiza comparações para determinar a similaridade entre a frase digitada pelo usuário e as frases presentes nos seus dados de treinamento. Com base nesse percentual podemos definir se uma frase se adequa a uma das presentes no nosso conjunto de informações teste e assim, recebe uma resposta. Portanto, podemos definir o algoritmo de distância de Levenshtein como sendo uma medida de distância de *string* que quantifica a distância entre as pronúncias das palavras correspondentes em diferentes dialetos ou línguas intimamente relacionadas. Calcula os custos mínimos necessária para transformar uma sequência de segmentos em outra por meio de inserções, exclusões ou substituições (Beijering, Gooskens, & Heeringa, 2008). O pseudocódigo deste algoritmo está representado na figura 3.

```
Função LevenshteinDistance(Caracter : str1[1..lenStr1], Caracter : str2[1..lenStr2]) : INTEIRO  
Início  
  // tab é uma tabela com lenStr1+1 linhas e lenStr2+1 colunas  
  Inteiro: tab[0..lenStr1, 0..lenStr2]  
  // X e Y são usados para iterar str1 e str2  
  Inteiro: X, Y, cost  
  
  Para X de 0 até lenStr1  
    tab[X, 0] ← X  
  Para Y de 0 até lenStr2  
    tab[0, Y] ← Y  
  
  Para X de 1 até lenStr1  
    Para Y de 1 até lenStr2  
      Se str1[X] = str2[Y] Então cost ← 0  
      Se-Não cost ← 1 // Custo da substituição deve ser 1, deleção e inserção  
      tab[X, Y] := menor(  
        tab[X-1, Y] + 1, // Deletar  
        tab[X, Y-1] + 1, // Inserir  
        tab[X-1, Y-1] + cost // Substituir  
      )  
  LevenshteinDistance ← tab[lenStr1, lenStr2]  
Fim
```

Figura 3: Psuedocódigo Distância de Levenshtein.

Outrossim, como o chatbot utiliza de técnicas de aprendizado de máquina, este então é capaz de reconhecer respostas para determinadas frases a medida que o chatbot é usado e sofre da aquisição de informações. Dessa forma, para decidir uma resposta para a frase digitada pelo usuário, é usado o algoritmo *Best Match* presente na biblioteca chatterbot. Logo, quando duas sentenças atingem o mínimo de similaridade necessária, medidas no algoritmo de Levenshtein, para serem consideradas pares, cabe ao algoritmo *Best Match* selecionar qual será a resposta enviada ao usuário. Assim este algoritmo simples seleciona a sentença presente em sua biblioteca que teve maior semelhança e esta é escolhida como resposta.

Vale ressaltar ainda que um pequeno trabalho de pré-processamento é feito com a frase submetida pelo usuário no momento de comparação para ajudar na compreensão das sentenças. Inicialmente a remoção de espaços em branco consecutivos é realizada e, além disso, a conversão de caracteres para os caracteres equivalentes e presentes na tabela ASCII. Por conseguinte, uma frase como: “Gosto de mineração” se transforma em “Gosto de mineracao”.

Para mais, outro aspecto que deve ser destacado é no uso da gamificação que, apesar de não ter um enfoque neste projeto, ainda sim pode-se ser aplicada dado que o chatbot oferece uma forma de diálogo diferente e que instiga o aluno a conversar com este a fim de conhecer essa tecnologia e além disso, procurar direcionamento por meio de um software desenvolvido em sua própria universidade. Assim, essa ferramenta ganha um tipo de atenção diferenciada quando comparadas com outras ferramentas de suporte convencionais no quesito de suporte e incitar o

aluno a conhecer melhor seu direcionamento dentro dos cursos de computação.

4 Análise Comparativa

Para realizar um estudo de viabilidade da proposta perante às outras técnicas de gamificação presentes e já exploradas em outros trabalhos científicos, um trabalho com o intuito de aferir os campos cobertos e diferenciais da proposta aqui apresentada quando comparada com outras pesquisas presentes na literatura foi feito. Isto posto, na figura 4 é observada uma tabela comparativa entre os trabalhos que serviram como base no desenvolvimento da proposta com a finalidade de salientar as lacunas preenchidas pelo projeto de pesquisa aqui feito.

	Chatbot	Usado para fins educacionais	Capaz de recomendar materiais	Open Source	Ajuda no direcionamento do aluno
<i>Chatbot for education system</i>	✓	✓	✗	✗	✗
<i>Chatbot: An Education Support System for Student</i>	✓	✓	✗	✗	✗
<i>Conversational Support for Education</i>	✓	✓	✓	✗	✗
<i>Doly: Bengali Chatbot for Bengali Education</i>	✓	✓	✗	✗	✗
<i>Personalized recommendation based collaborative filtering</i>	✗	✓	✓	✗	✗
<i>Turing</i>	✓	✓	✓	✓	✓

Figura 4: Comparação de diferentes tipos de *trabalhos*.

Com a figura é possível perceber os campos que foram primordiais no desenvolvimento da ideia e implementação desta tecnologia. Logo, o uso da tecnologia de chatbot para fins educacionais que fosse capaz de recomendar materiais e assistir no direcionamento do aluno foram questões de suma importância além de um projeto completamente *Open Source* para que assim

a comunidade acadêmica da universidade pudesse realizar melhoramentos conforme interesse e ache necessário. Dessarte, durante o processo de busca por trabalhos relacionados buscamos assimilar com diversos trabalhos o panorama acadêmico em cada um dos campos que julgamos essenciais e assim, com cada técnica e informação, construímos o aprendizado necessário para o progresso deste trabalho.

4.1 Considerações Finais

References

- Akcora, D. E., Belli, A., Berardi, M., Casola, S., Di Blas, N., Falletta, S., ... others (2018). Conversational support for education.
- Beijering, K., Gooskens, C., & Heeringa, W. (2008). Predicting intelligibility and perceived linguistic distance by means of the levenshtein algorithm.
- Busarello, R. I. (2016). *Gamification: princípios e estratégias*.
- Clarizia, F., Colace, F., Lombardi, M., Pascale, F., & Santaniello, D. (2018). Chatbot: An education support system for student.
- Drachler, H., Verbert, K., Santos, O. C., & Manouselis, N. (2015). Panorama of recommender systems to support learning. In *Recommender systems handbook* (pp. 421–451).
- Ebert, C., & Duarte, C. H. C. (2018). Digital transformation.
- Hiremath, G., Hajare, A., Bhosale, P., Nanaware, R., & Wagh, K. (2018). Chatbot for education system.
- Huang, W. H.-Y., & Soman, D. (2013). Gamification of education.
- Kowsher, M., Tithi, F. S., Alam, M. A., Huda, M. N., Moheuddin, M. M., & Rosul, M. G. (2019). Doly: Bengali chatbot for bengali education.
- Nah, F. F.-H., Zeng, Q., Telaprolu, V. R., Ayyappa, A. P., & Eschenbrenner, B. (2014). Gamification of education: a review of literature.
- Ortiz Rojas, M. E., Chiluiza, K., & Valcke, M. (2017). Gamification in computer programming: Effects on learning, engagement, self-efficacy and intrinsic motivation.
- Salehi, M., Kamalabadi, I. N., & Ghouschi, M. B. G. (2014). Personalized recommendation of learning material using sequential pattern mining and attribute based collaborative filtering.