

UM SISTEMA DE RECOMENDAÇÃO ENTRE ESTUDANTES E REPÚBLICAS BASEADO NA PERSONALIDADE

Marcos Paulo Ferreira Rodrigues

Departamento de Computação
Instituto de Ciências Exatas e Biológicas
Universidade Federal de Ouro Preto

12 de Dezembro de 2019



- 1 Introdução
- 2 Revisão da Literatura
- 3 Desenvolvimento
- 4 Conclusão

Contextualização

- Calouro: Estudante que acabou de ingressar na Universidade
- República: Casa onde reside um grupo de estudantes.
- Na tomada de decisão do Calouro sobre qual República escolher, não é considerado a personalidade
 - Incompatibilidade entre personalidades de moradores de república pode gerar a saída de um ou mais moradores.

Justificativa

- Redução do problema de evasão nas repúblicas
 - Busca de empresas por profissionais que defendam específicos ideais sociais.
- Solução: recomendar corretamente o perfil do usuário que se encontra na busca desejada.

Objetivo Geral e Objetivos Específico

Objetivo Geral

A concepção e o desenvolvimento de um sistema de recomendação de pessoas a grupos aplicado no contexto de estudantes e repúblicas..

Objetivos Específicos

- Investigar formas de detectar e representar personalidades
- Desenvolver uma estratégia de indicar a personalidade de um grupo por meio das suas personalidades individuais;
- Desenvolver e avaliar uma estratégia de recomendação de pessoas à grupos.
- Implementar um sistema que identifique a personalidade de pessoas e grupos e possibilite recomendações

Perfil do Estudante

- Modelo Big Five
- Teste de Personalidade BFI

Perfil da República

- Personalidade definida de acordo com o perfil dos estudantes vinculados

Recomendação

- Lista de usuários/grupos
- Similaridade entre as personalidades

Fundamentação Teórica

Sistemas de Recomendação

- Filtros de informação para apresentar itens ou objetos
- Principais componentes: **Cliente** (*usuário*) e **Produto** (*item*)
- Produto: conteúdo, arquivo, informação, pessoa, objeto.
- Recomendação: função de mapeamento dos interesse dos clientes para obtenção de um ou mais produtos.
- Caracterizado de acordo com três eixos:
 - Tipos de entrada e saída
 - Entrada: histórico de compras, avaliações deixadas por usuários, extração de palavras-chave em consultas...
 - Saída: Nome, informação, página web, item...
 - Itens do projeto
 - Método de recomendação

Filtragem colaborativa

- Prevê o grau de interesse de um cliente em determinados produtos
 - Correlaciona as avaliações feitas pelo cliente com outros clientes.
- Método:
 - 1 Calcular a similaridade entre os usuários
 - 2 Selecionar os vizinhos mais próximos
 - 3 Fazer a previsão do usuário sobre um produto não consumido
- Vantagens:
 - Possibilita apresentar recomendações inesperadas ao usuário
 - Auxilia na formação de grupos de usuários agrupados por gostos e interesses similares
- Desvantagem:
 - Computacionalmente caro
 - Poucos usuários e avaliações, método tende a ser ruim

Filtragem baseada em conteúdo

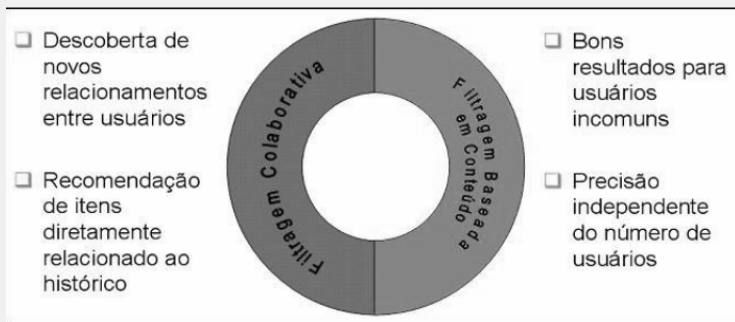
- Sugestão é feita de acordo com o que o usuário demonstrou interesse no passado.
 - Personalizada para cada usuário
- Método
 - 1 Classificar itens de acordo com categorias pré-estabelecidas
 - 2 Para cada categoria, calcular a avaliação média de cada avaliador
 - 3 Obter os itens avaliados
 - 4 Calcular a previsão de avaliação
- Vantagens:
 - Não precisa de dados de outros usuários
 - Permite recomendações a usuários com gostos exclusivos e itens novos/não populares.
- Desvantagens:
 - Recomendações desprezam opiniões dos usuários
 - Conteúdo precisa ter características inteligíveis
 - Baixa eficiência com conteúdo pouco informativo.

Filtragem baseada em Descoberta de Conhecimento

- Correspondências entre as necessidades do usuário e o conjunto de dados disponíveis
 - Recomendação baseada na frequência com que as ações acontecem conjuntamente
- Utiliza dois critérios: suporte e confiança
 - Suporte: frequência com que dois eventos aconteceram simultaneamente
 - Confiança: probabilidade de um evento acontecer, uma vez que outro evento aconteceu
- Vantagem:
 - Aumentar a precisão da recomendação e evitar limitações intrínsecas das duas abordagens citadas
- Desvantagem:
 - Requer "adquirir conhecimento- gargalo de diversas aplicações na área de Inteligência Artificial

Filtragem Híbrida

- Combinar os pontos fortes da filtragem baseada em conteúdo e colaborativa
- Unificação das duas técnicas
 - Utiliza as vantagens proporcionada pelas duas técnicas
 - Descarta suas desvantagens



Perfil do Usuário

- Antes de recomendar produtos, serviços ou pessoas a um usuário, é necessário ter conhecimento sobre quem é este usuário
- Identidade Visual
 - Identidade Social
 - Reputação
 - Definida pelos outros usuários
 - Identidade Interna
 - Definida pelo próprio usuário
 - Técnicas de machine learning

Coleta de Informações

- É necessário identificar o usuário no momento que ele acessa o sistema
- Identificação no servidor:
 - Disponibiliza um formulário
 - Solicita login/senha
 - Dados armazenados em Banco de Dados
- Identificação no cliente:
 - Utiliza cookies
 - Assume que cada computador é usado por uma pessoa
- Após identificar o usuário, a coleta de informações pode ser implícita ou explícita
 - Explícita: cliente indica de forma espontânea os seus interesses
 - Implícita: informações sobre necessidades e preferências do usuário são adquiridas de acordo com suas ações

Traço de Personalidade

Personalidade:

- Relativamente estáveis e previsíveis
- Não são rígidas ou imutáveis
- Podem ser definidas por diferentes abordagens

Traços:

- Padrões habituais de comportamento, pensamento e sentimento
- Foram definidos como 17.953 traços
- Somente cinco são replicáveis
- Nasce o Big Five Model

Como obter os traços de personalidade?

- Testes de Personalidade
 - NEO-PI-R (240 perguntas), NEO-IPIP (300 perguntas)
 - Big Five Inventory (44 perguntas)

Modelo Big Five

- Descreve as características humanas em cinco principais traços: neuroticismo, extroversão, socialização, realização e abertura para o novo. A combinação dos cinco fatores reflete os principais aspectos da personalidade humana
- **Abertura para o novo:** apreciação de arte, emoção, aventura, ideias incomuns, curiosidade e variedade de experiência
- **Realização:** tendência a mostrar autodisciplina, agir literalmente e realizar objetivos. Rígido em vez de um comportamento espontâneo.
- **Extroversão:** energia, emoções positivas, urgência e uma tendência a estimular o ambiente na companhia de outras pessoas.
- **Socialização:** tendência para ser mais compassivo e cooperativo do que desconfiado e antagonista em relação aos outros
- **Neuroticismo:** tendência a ter emoções negativas facilmente, como raiva, ansiedade, depressão e vulnerabilidade.

Teste de Personalidade

- Ferramenta computacional capaz de medir as diferenças individuais das pessoas
- Big Five Inventory
 - Desenvolvido pelo mesmo autor do Modelo Big Five.
 - Um dos inventários mais utilizados na psicologia
 - Inventário de autorrelato projetado para medir os cinco fatores do Modelo Big Five
 - BFI Original: 44 perguntas
 - Versão curta do BFI: 10 perguntas
 - Pode ser usado somente para fins acadêmicos
 - Tempo médio de 5 minutos para responder todo o inventário

Sistemas de Recomendação com Traços de Personalidade

- Proposto em 2008, por Nunes
- Utiliza Traços de Personalidade na Reputação Psicológica do Perfil do Usuário
- Reputação Psicológica é o Perfil Psicológico do Usuário de acordo com a visão de seus amigos
- Caso de Estudo
 - Eleições para presidente na França - 2006
 - Inventário NEO-IPIP
 - Experimentou durou de Dez 2006 até Jul 2007
 - Dois tipos de recomendações: Utilizando 6 facetas e utilizando somente o Big Five.
 - Resultado com 6 facetas: Precisão de 100%
 - Resultado somente com Big Five: Precisão de 80%

Personalidade e Confiança Social em Grupos de Recomendação

- Maximiza a satisfação do grupo usando a personalidade das pessoas e relações sociais entre pessoas do grupo
- Três sistemas de recomendação: recomendação básica; recomendação considerando personalidade; recomendação considerando personalidade e confiança social;
- Caso de Estudo
 - Recomendação de filme para um grupo
 - Recomendação se baseia na mistura de recomendações individuais.
 - Recomendação individual é um sistema colaborativo baseado na avaliação de outros usuários
 - Utiliza o Teste de Personalidade TKI
 - Resultado: a utilização de confiança social em sistemas de recomendação influencia na tomada de decisão de um grupo

Prever Traços de Personalidade com Imagens do Instagram

- Propõe um novo modelo de estimar as preferências do usuário sem usar questionários ou observações
- Utiliza o Modelo Big Five
- Utiliza características da imagem para criar uma correlação entre o usuário e sua personalidade
- Resultados sugerem a existência de uma relação entre traços de personalidade e modo como usuários publicam fotos
- Resultados satisfatórios para: Abertura a experiência, Extroversão e Neuroticismo.

Sistema de Recomendação

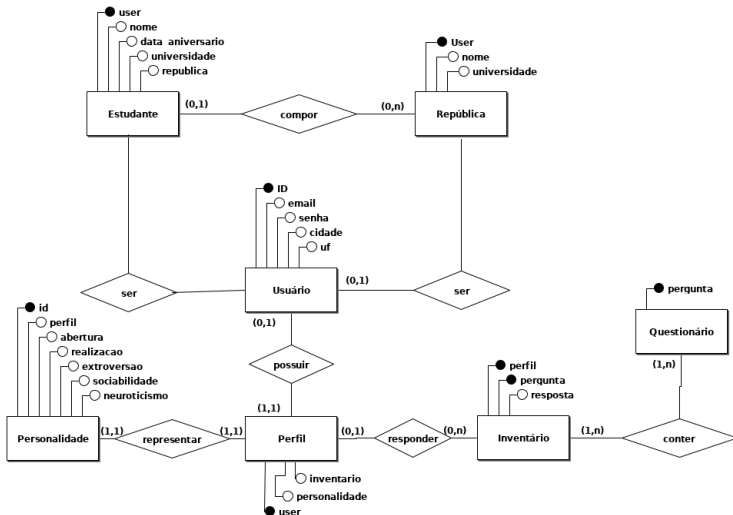
- Recomendar, de maneira adequada, um estudante a uma república, baseado na similaridade entre as personalidades.
- Ferramenta acessada via *Web*
- Perfil do Usuário
 - Identificação no Servidor
 - Modelo de Personalidade Big Five
 - Teste de Personalidade BFI
- Perfil da República
 - Formada de acordo com a Personalidade dos usuários que compõem o grupo
- Recomendação baseada em distância das personalidades

$$D_{u_i, u_j} = \sum_{k=1}^{k=5} |f_{k_i} - f_{k_j}| \quad (1)$$

Lista de relações no banco de dados relacional

- Usuário (**id**, email, senha, cidade, uf)
- Estudante (**user**, nome, birth, universidade, republica)
- Republica (**user**, nome, universidade, moradores)
- Perfil (**user**, personalidade, inventario)
- Personalidade (**id**, perfil, abertura, realizacao, extroversao, sociabilidade, neuroticismo)
- Questionario (**pergunta**)
- Inventario (**user**, pergunta, resposta)

Desenvolvimento - Diagrama Entidade Relacionamento



Concepção do Sistema

- Objetivo
- Descrição
- Missão
- Escopo
- Diagramas
- Atores
- Storyboard
- Backlog
- Estórias de Usuários e Casos de Teste

Objetivo

Desenvolver um sistema *Web* capaz de recomendar ao estudante as repúblicas que mais possuem compatibilidade com seu perfil

Descrição

Nome e Componentes

- Find My Republic - FMR
- Dois componentes: dependem do ator utilizando o sistema (Estudante ou República)

Descrição

- Sistema *Web* que oferece serviços de relatórios de personalidade dos estudantes e repúblicas obtidos a partir do BFI

Missão

- Auxiliar na evasão de estudantes nas repúblicas
- Contribuir para que o estudante encontre uma república com o perfil desejado
- Ajudar repúblicas a encontrar estudantes que complementem o perfil da república

Escopo

- Positivo: estudante pode receber uma recomendação adequada.
- Negativo: estudante pode receber uma recomendação inadequada.

Diagrama de Contexto

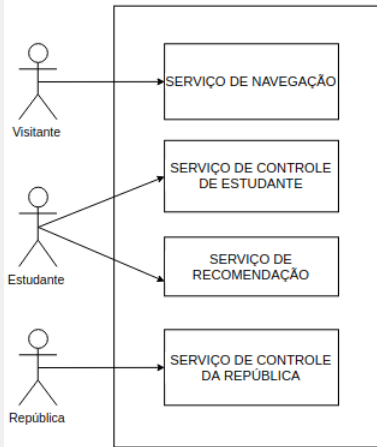
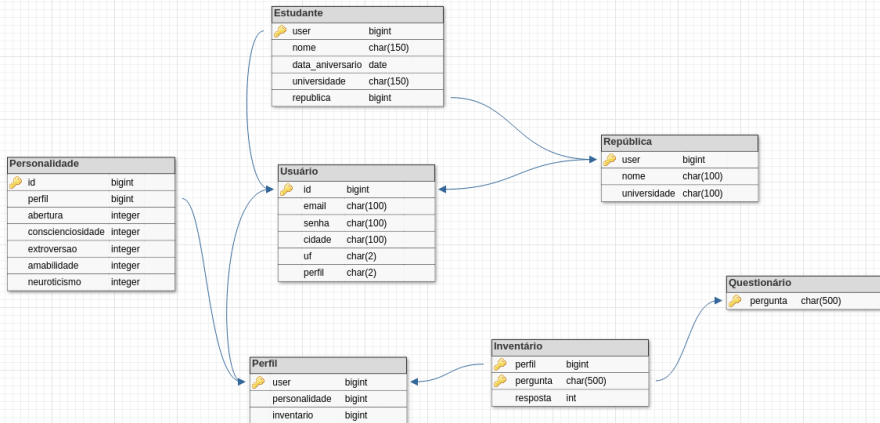
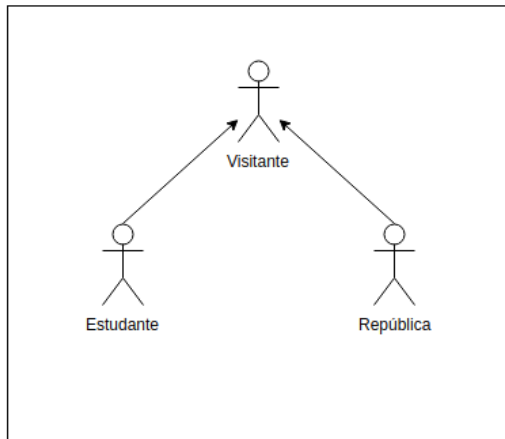


Diagrama Relacional



Hierarquia dos Atores



Serviço de Navegação

- 1 Efetuar consulta
- 2 Acessar república
- 3 Efetuar cadastro

Serviço de Controle de Estudante

- 1 Efetuar login
- 2 Alterar dados
- 3 Visualizar dados
- 4 Excluir conta
- 5 Atualizar personalidade
- 6 Vincular a uma república

Serviço de Recomendação

- 1 Executar recomendação

Serviço de Controle de República

- 1 Efetuar login
- 2 Alterar dados
- 3 Visualizar dados
- 4 Excluir conta
- 5 Atualizar personalidade
- 6 Vincular estudante

Storyboards

- Execução do sistema de recomendação.

Estórias de Usuário e Cenários de Teste

- Tais seções não serão apresentadas devido a falta de tempo, visto que estórias e cenários de testes são longos
- A descrição e detalhamento destas seções estão descritas no projeto da Monografia

Tecnologias

- Angular 8.1.3v; Node 11.6v; Bootstrap 4.0v
- Java 8; SpringBoot 2.1.7v; MySQL 5.7v
- Google Cloud App Engine; Google Cloud Storage; Google Cloud SQL

Descrição

- Sistema de recomendação web que fornece o serviço de recomendar repúblicas à estudantes de acordo com a personalidade.
- Componentes: Estudante e República
 - Estudante: Perfil do Usuário
 - República: Perfil do Grupo

Personalidade

- Big Five. Abertura para o novo: [10,50]; Concordância e Consciência: [9,45]; Extroversão e Neuroticismo: [8,40]

Personalidade Normalizada

Obtida de acordo com a pontuação de cada categoria da personalidade

$$PN_{u_i} = \frac{P_{u_i}}{\max(P_i)} * 100\% \quad (2)$$

- PN_{u_i} : personalidade normalizada do usuário u sobre a faceta i
- P_{u_i} : personalidade obtida no teste pelo usuário u sobre a faceta i
- $\max(P_i)$: máximo valor que pode ser obtido da categoria i

Perfil do Usuário

- Big Five Inventory

Perfil da República

- Agregação da personalidade dos moradores da seguinte maneira:
 - 1 Obtêm-se as Personalidades ps dos estudantes vinculados à república
 - 2 Para cada estudante u , categoriza cada Faceta i da Personalidade p seguindo a regra:
 - 1 Calcula a Personalidade Normalizada PN da faceta i
 - 2 Obtém a Categoria C da norma personalidade do usuário u para a faceta i , sendo:

$$C_{u_i} = \left\lceil \frac{PN_{u_i}}{20} \right\rceil \quad (3)$$

- 3 Calcula as modas das Categorias c de todos os estudantes u com base em cada faceta i .
- 4 Obtém a maior moda para cada faceta i .
- 5 Calcula a média das personalidades dos estudantes para a faceta correspondente à categoria.
- 6 A Personalidade P para a faceta i é resultado da média da faceta i de todos os estudantes u que foram obtidos através da maior moda das categorias para a faceta i .

Perfil da República

$$PR_i = \frac{\sum_{u=1}^N \alpha_u * P_{u_i}}{|\max(modas(C_i))|}, \alpha_k = \begin{cases} 1, & \text{se } P_{u_i} \in \max(modas(C_i)) \\ 0, & \text{caso contrário} \end{cases} \quad (4)$$

Sendo:

- PR_i : Personalidade da República sobre a Faceta i
- N : número de estudantes vinculados à República
- P_{u_i} : Personalidade do estudante u sobre a faceta i .
- C_i : Vetor de categoria com as personalidades dos estudantes categorizadas correspondente a faceta i
- $modas(C_i)$: lista de personalidades dos estudantes que compõem as modas da categoria i .
- $\max(modas(C_i))$: lista de estudantes que foram obtidos a partir da maior moda da categoria C para a faceta i .

Recomendação

- Recomendação é calculada de acordo com a distância das personalidades

$$R_{u,v} = |D_{u,v}| \quad (5)$$

sendo:

- $R_{u,v}$: recomendação do usuário u para o usuário v
- $D_{u,v}$: distância definida pela fórmula:

$$D_{u,v} = 1 - |Eq_{u,v}| \quad (6)$$

sendo:

$$Eq_{u,v} = \begin{cases} \frac{\sum_{i=1}^5 P_{v,i}}{\sum_{i=1}^5 P_{u,i}}, & \sum_{i=1}^5 P_{u,i} \geq \sum_{i=1}^5 P_{v,i} \\ \frac{\sum_{i=1}^5 P_{u,i}}{\sum_{i=1}^5 P_{v,i}}, & \text{caso contrário} \end{cases} \quad (7)$$

Descrição

- Dados reais, ou seja, após a extração de dados realizada pelo sistema de recomendação.
 - 1 Obtenção da personalidade dos estudantes
 - 2 Atribuição de vínculos entre estudantes e repúblicas
 - 3 Calcular a personalidade das personalidades
 - 4 Aplicar separação entre estudantes e repúblicas
 - 5 Aplicar a recomendação
 - 6 Analisar recomendação
- Experimentos podem predizer algumas situações:
 - Proximidade dos estudantes às repúblicas
 - Eficácia da recomendação nos critérios analisados

Critério de avaliação do Estudante

- Estudante vinculado à república
 - República do estudante pertence às 3 primeiras repúblicas vinculadas ou
 - República do estudante tenha compatibilidade igual ou superior a 80%
- Estudante não vinculado à república
 - Existe vínculo de amizade entre estudante e república
 - Repúblicas recomendadas estejam entre as repúblicas amigas
- República solicitando recomendação:
 - Semelhante ao caso de Estudante não vinculado à república

Dados gerais

- 120 usuários: 23 repúblicas e 97 estudantes;
- 97.96% residem em Ouro Preto ou Mariana
- Usuários válidos: 62 personalidades. 11 repúblicas e 51 estudantes. 11 estudantes desvinculados e 40 estudantes vinculados.

Estudantes sem vínculo com República

- Aplicada a 11 estudantes sem vínculos com república
- Considerado a existência de amizade
 - Coletada de forma direta, verbalmente.

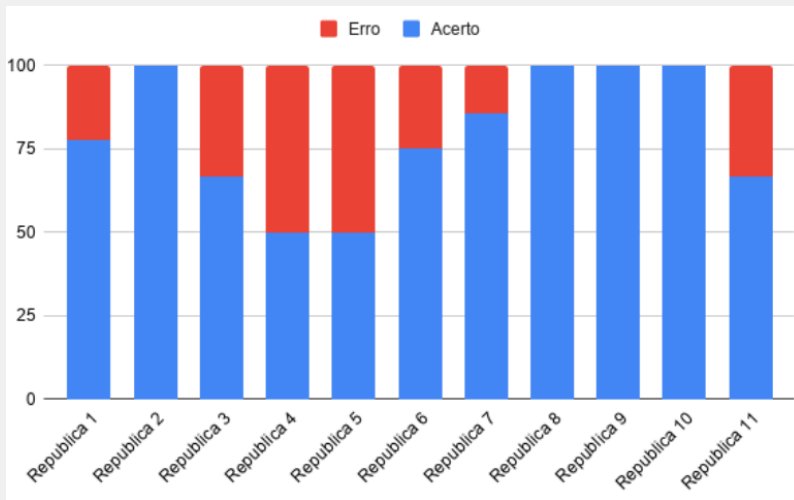
Estudante	Nº Rep Amigas	Nº Rep Não Amigas	Precisão
Estudante 1	4	1	80%
Estudante 2	5	0	100%
Estudante 3	4	1	80%
Estudante 4	5	0	100%
Estudante 5	5	0	100%
Estudante 6	4	1	80%
Estudante 7	3	2	60%
Estudante 8	2	3	40%
Estudante 9	4	1	80%
Estudante 10	2	3	40%
Estudante 11	4	1	80%
Total	42	13	76.36%

Estudantes moradores de República

- Aplicada a 51 estudantes moradores de alguma república
- República do estudante deve estar entre as 3 primeiras ou compatibilidade igual ou superior a 80%.
- Resultado de cada república é a média dos resultados de seus moradores

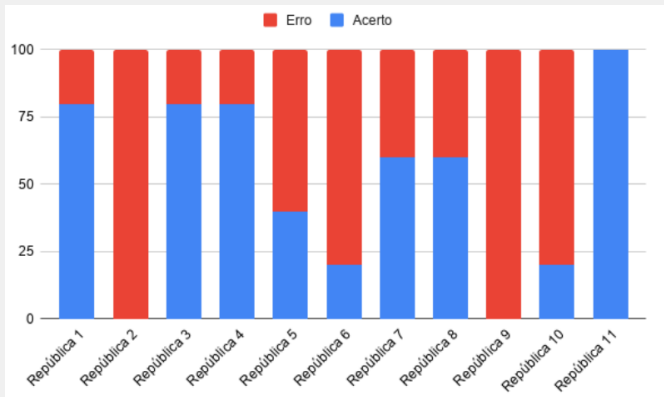
República	Nº de moradores
República 1	9
República 2	2
República 3	3
República 4	2
República 5	2
República 6	4
República 7	7
República 8	2
República 9	3
República 10	4
República 11	3
Total	41

Estudantes moradores de República



Recomendação aplicada à Repúblicas

- 11 repúblicas
- Considerado a existência de amizade
 - Coletada de forma direta, verbalmente



Estudantes sem vínculo com república

- Objetivo: obter repúblicas amigas como primeiras repúblicas sugeridas.
- Acurácia de 76,36%.

Estudantes moradores

- Objetivo: obter como sugestão a república onde o estudante é morador
- Repúblicas com 4 ou menos moradores, dois cenários:
 - Estudantes que compõem a república possuem uma diferença significativa em suas personalidades, conduzindo a o perfil da república a ter uma personalidade versátil. Acurácia entre 50% e 75%.
 - Estudantes que compõem a república são próximos, conduzindo o perfil da república a ter uma personalidade mais consistente. Acurácia de 100%.
- Repúblicas com 7 ou mais moradores: personalidade consistente, recomendações mais eficazes. Acurácia entre 80% e 90%.

Conclusão

Observações gerais

- Sistemas de recomendação podem contribuir para o problema da evasão das repúblicas
- Testes de Personalidade podem contribuir para a acurácia das recomendações baseadas em traços de personalidade

Conclusão geral

- Foram abordadas as etapas que comõem o desenvolvimento de um sistema de recomendação baseado em traços de personalidade
- Foi apresentado uma pesquisa na literatura para obter a fundamentação teórica e os trabalhos relacionados ao tema desta monografia
- Foi definido o escopo e a documentação do sistema.
- Foram apresentados os experimentos computacionais e seus resultados

Principais resultados

- Ferramenta de recomendação aplicável e pode auxiliar no problema de evasão de repúblicas
- Para recomendações mais eficientes é recomendado que as repúblicas possuam mais que 04 moradores, uma vez que a personalidade da república pode ficar versátil.

Trabalhos futuros

- Aplicar o sistema de recomendação a calouros que estejam procurando por uma república
- Possibilitar repúblicas a procurar por estudantes
- Possibilitar avaliação explícita sobre as recomendações recebidas

Dúvidas?