

Sistema de Simulados com Recomendações baseado no Desempenho do Usuário

**Trabalho de Conclusão do Curso de
Tecnologia em Sistemas para Internet**

**Marcos Henrique Zotti de Sá
Orientadora: Márcia Hafele Islabão Franco**

¹Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS)
Campus Porto Alegre
Av Cel Vicente, 281, Porto Alegre – RS – Brasil

marcoshzotti@gmail.com, marcia.franco@poa.ifrs.edu.br

***Resumo.** Atualmente existe uma grande quantidade de sites de simulados, tanto para vestibulares quanto para concursos, porém, observa-se a falta de aplicação de técnicas inovadoras, visando uma atenção maior ao usuário. Neste contexto, este artigo descreve o desenvolvimento de um sistema web. O sistema tem como objetivo auxiliar os usuários durante o desenvolvimento das atividades através de recomendações. As recomendações são sugeridas com base no desempenho do usuário e conta com a aplicação da técnica de mineração de dados. No desenvolvimento deste trabalho, foram pesquisados três sistemas semelhantes ao sistema proposto, o que possibilitou realizar uma comparação entre os sistemas, destacando as vantagens do proposto, e o levantamento das funcionalidades desejadas.*

1. Introdução

Existe uma grande procura por cursos online que têm por objetivo disponibilizar simulados que são utilizados por usuários para se prepararem para concursos e vestibulares. Porém percebe-se que eles apresentam muitas semelhanças, como as funcionalidades e o método de resolução de provas. Dessa forma, observa-se a falta de aplicação de técnicas inovadoras, visando uma atenção maior ao usuário.

Neste sentido, o trabalho proposto tem por objetivo o desenvolvimento de um sistema de simulados desenvolvido através do uso das técnicas de Sistemas de recomendação e Mineração de Dados. O sistema auxilia o usuário durante o desenvolvimento das atividades, fazendo com que ele responda questões mais fáceis, caso seja necessário, e posteriormente responder questões mais difíceis, até atingir um bom desempenho, que neste caso, pode ser considerado uma nota superior a nota de corte.

O sistema recomenda ao usuário as atividades que ele pode realizar, sendo essas recomendações sugeridas de acordo com o desempenho dele durante as atividades. Para realizar as recomendações, o sistema utiliza o método de classificação da mineração de dados.

O artigo está organizado da seguinte forma: na seção 2 são apresentados os conceitos envolvidos no desenvolvimento do trabalho, os trabalhos relacionados são discutidos na seção 3, o sistema é apresentada na seção 4, os resultados são apresentados na seção

5, a metodologia utilizada no trabalho é descrita na seção 6. Por fim, na seção 7 são abordadas as considerações finais.

2. Conceitos Envolvidos

Nesta seção serão abordados os conceitos utilizados no desenvolvimento do trabalho.

2.1. Sistemas de Recomendação

Os sistemas de recomendação auxiliam o usuário na busca de produtos e informações desejadas, baseado em um histórico de preferências. São muito utilizados como uma forma de marketing, pois assim, facilitam a venda de itens quando já são conhecidos os interesses do usuário [Cazella 2011]. Para adquirir essas informações que serão utilizadas para a recomendação, os dados do usuário são obtidos normalmente de duas formas: explícita ou implícita.

A forma explícita é a forma direta com o usuário de adquirir informações, através de formulários, por exemplo, que contém questões sobre dados pessoais e preferências. Na forma implícita, os dados são coletados a partir do monitoramento da navegação do usuário em um site ou pelo histórico. As informações então são armazenadas em um banco de dados para que possam ser analisadas para uma recomendação futura [Cazella 2011].

Os sistemas de recomendação podem ser classificados em três tipos [Cazella 2011]:

- Baseado em Conteúdo - O sistema recomenda produtos com características semelhantes aos que um determinado usuário visualizou ou adquiriu no passado. É um sistema simples, não necessita de muitas informações além das preferências do usuário. Porém, ele sofre do problema de muitas vezes oferecer recomendações repetitivas, trazendo sempre algo que o usuário já conhece.
- Filtragem Colaborativa - Utiliza a técnica de recomendar itens com base em pessoas com gostos semelhantes ao usuário preferiram no passado. Esse tipo de sistema elimina o problema da recomendação repetitiva, mas ele requer um grande número de informações de usuários para funcionar com precisão.
- Sistemas Híbridos - Consistem em combinar os outros dois tipos, tentando fortificá-los e superando suas desvantagens. Um sistema híbrido utiliza também de técnicas de mineração de dados. Esse tipo de recomendação foi aplicada no trabalho e será abordada na seção 2.2.

Os websites de comércio eletrônico são atualmente os maiores utilizadores dos sistemas de recomendação [Cazella 2011], pois assim, encontram com mais facilidade o produto que seus clientes desejam, porém outros sites famosos por utilizarem de sistemas de recomendação são as redes sociais. O Facebook, por exemplo, utiliza dessas técnicas de recomendação com seus usuários para indicar quais pessoas podem ser suas amigas, baseados nos amigos atuais dela.

No sistema proposto, a técnica de recomendação será utilizada para auxiliar na escolha da questão que o usuário deverá realizar naquele determinado momento, com base no desempenho dele nas questões anteriores. Outros fatores também serão avaliados para que o sistema faça a recomendação, essas variáveis serão comentadas nas próximas seções.

2.2. Mineração de Dados

Mineração de Dados é um conceito que faz parte de uma área conhecida como KDD (*Knowledge Discovery in Databases*) – em português, Descoberta de Conhecimento em Bases de Dados, que é um processo que atua em cima de bases de dados [Sferra 2003]. O KDD é composto por diversas etapas, sendo a principal delas a mineração de dados.

Segundo [Fayyad 1996] as etapas do processo de KDD são: i) Seleção: consiste no estudo dos dados disponíveis e na importância desses para a busca de soluções para os problemas identificados; ii) Pré-processamento: consiste em ajustar os dados a normas pré-estabelecidas, uniformizando a forma de representação dos mesmos; iii) Transformação: consiste na transformação dos dados num formato apropriado para aplicação das técnicas de mineração de dados; iv) Mineração dos Dados: conforme anteriormente descrito, principal etapa do KDD, que consiste na aplicação das técnicas de mineração de dados aos dados previamente pré-processados; e v) Avaliação dos Resultados: consiste em analisar o conhecimento gerado e utilizá-lo para auxiliar na solução dos problemas que motivaram a mineração dos dados.

Existem diversos métodos que podem ser utilizados para minerar os dados, muitos deles baseados em técnicas das áreas de aprendizado de máquina, reconhecimento de padrões e estatística [Sferra 2003]. Porém, para todos esses métodos, é necessário que as metas e os objetivos da mineração de dados sejam bem definidos.

Uma das principais aplicações da mineração de dados é em Business Intelligence (BI), pois os sistemas de informação ajudam a promover a inovação quando geram grande volume de dados atuais, confiáveis, acessíveis e consistentes que promovem agilidade, comunicação e integração com outros sistemas [Coelho 2015]. O BI é utilizado para explorar as potencialidades de relacionamento que um sistema pode ter com um usuário.

No trabalho proposto, a mineração de dados será responsável por analisar os dados das provas realizadas, transformando esses dados em informação para que no futuro seja aplicada na recomendação das questões. Como as técnicas de mineração de dados são aplicadas em diversos sistemas de recomendação, neste trabalho não será diferente, a mineração implicará diretamente no sistema de recomendações.

3. Trabalhos Relacionados

Com a vastidão de sistemas sobre simulados, foi realizada uma pesquisa para o desenvolvimento deste trabalho onde foi possível encontrar, características semelhantes às do sistema proposto. Os trabalhos pesquisados foram: Trilha do ENEM, Gabarite e Casa das Questões.

Lançado em 2015, o "Trilha do ENEM" [Trilha do Enem 2015] é um sistema de resolução de simulados, focado principalmente nas provas do Exame Nacional do Ensino Médio. O Trilha do ENEM possui diversas opções para que o usuário faça seus simulados, como provas rápidas, por área ou personalizada. Após terminar o simulado, o usuário pode optar por salvar a prova para revê-la e refazê-la no futuro. Quando realizado um simulado, são exibidas informações e estatísticas do desempenho do usuário nas questões que foram respondidas. Além disso, a plataforma disponibiliza estudos de recomendação. Embora sejam feitas recomendações, elas não são implementadas com técnicas de mineração de dados e nem são aplicadas a características que um sistema de

recomendação deve apresentar.

Outro sistema existente é o "Gabarite"[Gabarite 2010], que contém mais de cinco mil simulados para serem resolvidos e está no ar desde 2010. Suas provas podem ser filtradas por nível de escolaridade, disciplinas, instituições, etc. Antes de realizar um simulado, é possível verificar a quantidade de usuários que já o completaram e como foi o seu desempenho. O sistema também disponibiliza resumos, vídeos e outros materiais didáticos.

Por fim, o último sistema desta pesquisa a ser apresentado é o "Casa das Questões"[Casa das Questões 2014], que está disponível desde 2014. O site de simulados conta com provas dos principais concursos do Brasil. Nele é possível escolher qual concurso se deseja simular uma prova, ou, se o usuário preferir, ele pode optar por criar uma prova customizada, selecionando quais matérias serão abordadas na prova. As provas podem ser realizadas do modo teste, onde é possível responder as questões e imediatamente ver o resultado delas e pode-se livremente corrigi-las e prosseguir com a prova. Ou então, pode se optar pelo modo simulado, que só permite visualizar os resultados após o término da prova, além de contar com um cronômetro que caso o limite de tempo chegar, a prova é encerrada.

Após a realização da prova, são exibidos gráficos e estatísticas do desempenho do usuário. Um detalhe interessante, que antes de realizar um simulado de algum concurso, é possível verificar quais matérias caem naquela prova e qual o peso que elas terão na nota. Embora o sistema seja gratuito, ele disponibiliza planos de assinatura, caso o usuário se interesse pelos materiais didáticos para estudo exclusivos que o sistema em questão possui. Além, também, de possibilitar que o assinante possa sugerir questões e concursos para o site.

A seguir, é exibido na Tabela 1, um comparativo dos sistemas pesquisados no decorrer deste trabalho e a comparação com o sistema proposto.

Tabela 1. Comparativo dos Sistemas

Sistemas/Funcionalidades	Trilha do ENEM	Gabarite	Casa das Questões	Solução proposta
Gratuito	X	X	X	X
Desempenho do Usuário	X		X	X
Revisão das Provas	X	X	X	X
Recomendação de questões adequadas ao usuário				X
Avaliação de Perfil				X
Utilização sem Cadastro				X

Para melhor compreender as características destacadas na Tabela 1, a seguir são detalhadas as funcionalidades:

1. Gratuito: É altamente importante que um sistema como esse seja gratuito e livre, para que possa abranger um grande número de usuários e auxiliar a todos. Por esse motivo é possível perceber que todos os sistemas relacionados são gratuitos.
2. Desempenho do Usuário: Todas as questões que forem respondidas devem ser armazenadas para que o usuário possa analisar seu desempenho e compare com a média dos outros que utilizam o sistema. Apenas o sistema Gabarite não exibe o desempenho completo do usuário.

3. Revisão das Provas: As provas que forem feitas devem ficar salvas para que no futuro elas possam ser revisadas e refeitas.
4. Avaliação de Perfil: O perfil do usuário é determinante para o tipo de questões que ele deve responder nos simulados.
5. Recomendação de Questões adequadas ao usuário: O desempenho do usuário pode ser elevado caso sejam recomendadas questões com base no próprio.
6. Utilização sem Cadastro: Utilizar o sistema sem ser necessário fazer um cadastro permite atender um maior número de usuários.

4. Sistema Desenvolvido

Conforme apresentado anteriormente, existem inúmeros sistemas de resolução de simulados online e a maioria deles podem ser considerados qualificados para que os seus usuários possam utilizá-los e, assim, obterem um bom desempenho em uma prova de vestibular ou concurso. Porém, percebe-se que sistemas semelhantes continuam a surgir, mas acabam não trazendo inovações para o tema, abordando sempre as mesmas características e mecânicas, conforme apresentado na seção 3.

Nesse sentido, esse trabalho apresenta o desenvolvimento de um sistema web de resolução de simulados, que adapta as questões baseado no desempenho do usuário, ou seja, se o utilizador estiver com dificuldade em acertar a maioria das perguntas, o sistema automaticamente passa a sugerir perguntas categorizadas mais fáceis, para que assim, o utente tenha um preparo maior para responder um item mais difícil. A recomendação das questões é através da utilização do método de classificação da técnica de mineração de dados.

Para iniciar as atividades, o usuário pode se cadastrar no sistema inserindo apenas alguns dados, ou ele pode optar por utilizar sem cadastro, mas sendo assim, não será possível revisar provas e questões feitas anteriormente, pelo fato de suas informações não ficarem salvas no sistema. Com o cadastro feito, o utilizador poderá iniciar seus simulados e verificar seu desempenho.

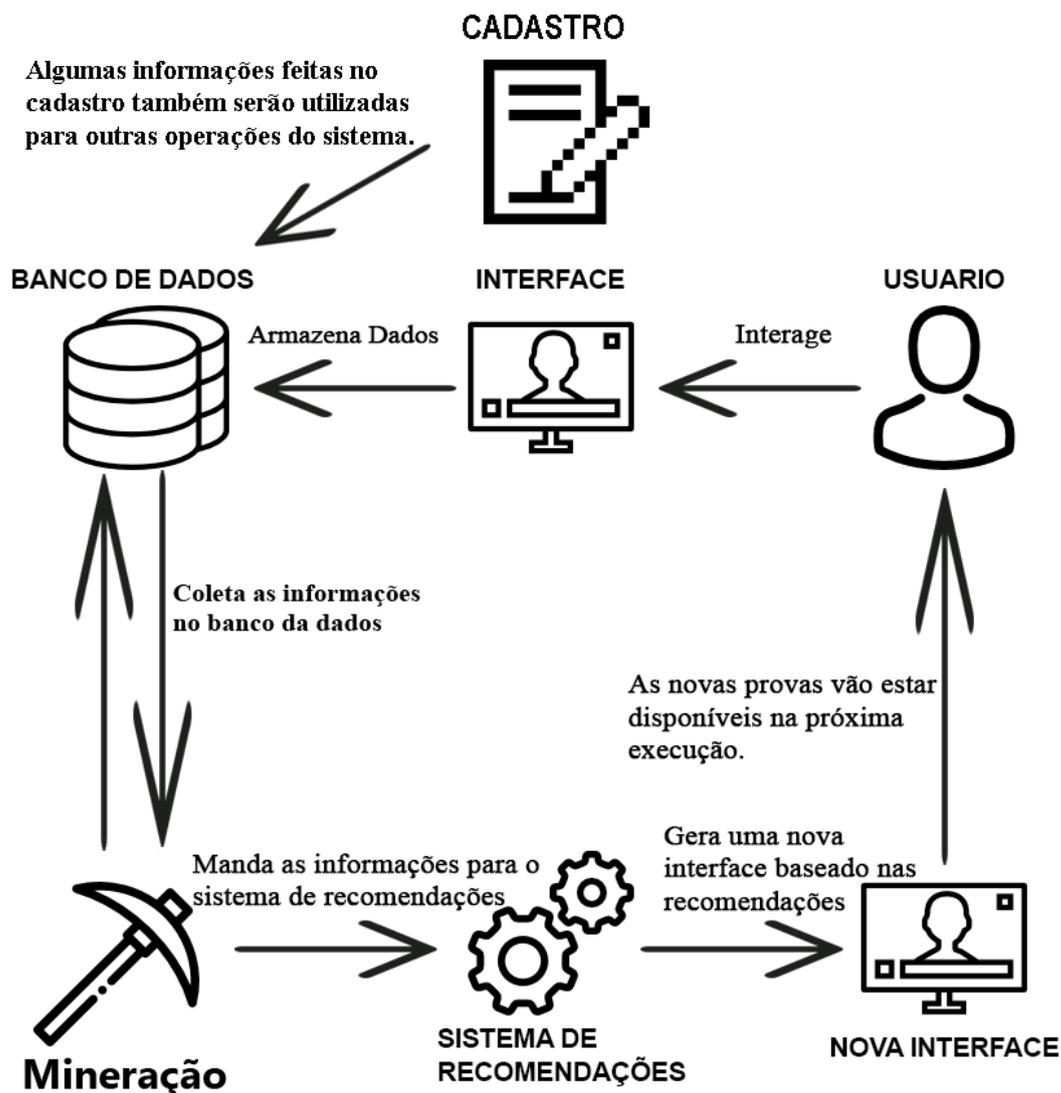


Figura 1. Diagrama de Visão Geral do Sistema.

A Figura 1 apresenta o diagrama de visão geral do sistema. O usuário do sistema ao realizar seus simulados, irá alimentar um banco de dados. Nesse banco será realizada a técnica de mineração de dados para coletar essas informações do usuário. Os dados então serão transformados pelo sistema de recomendação e salvos para a próxima vez que o usuário realizar um simulado.

A Figura 2 ilustra o diagrama de casos de uso [Franco 2016] do sistema proposto tanto do ator Usuário, quanto do Administrador.

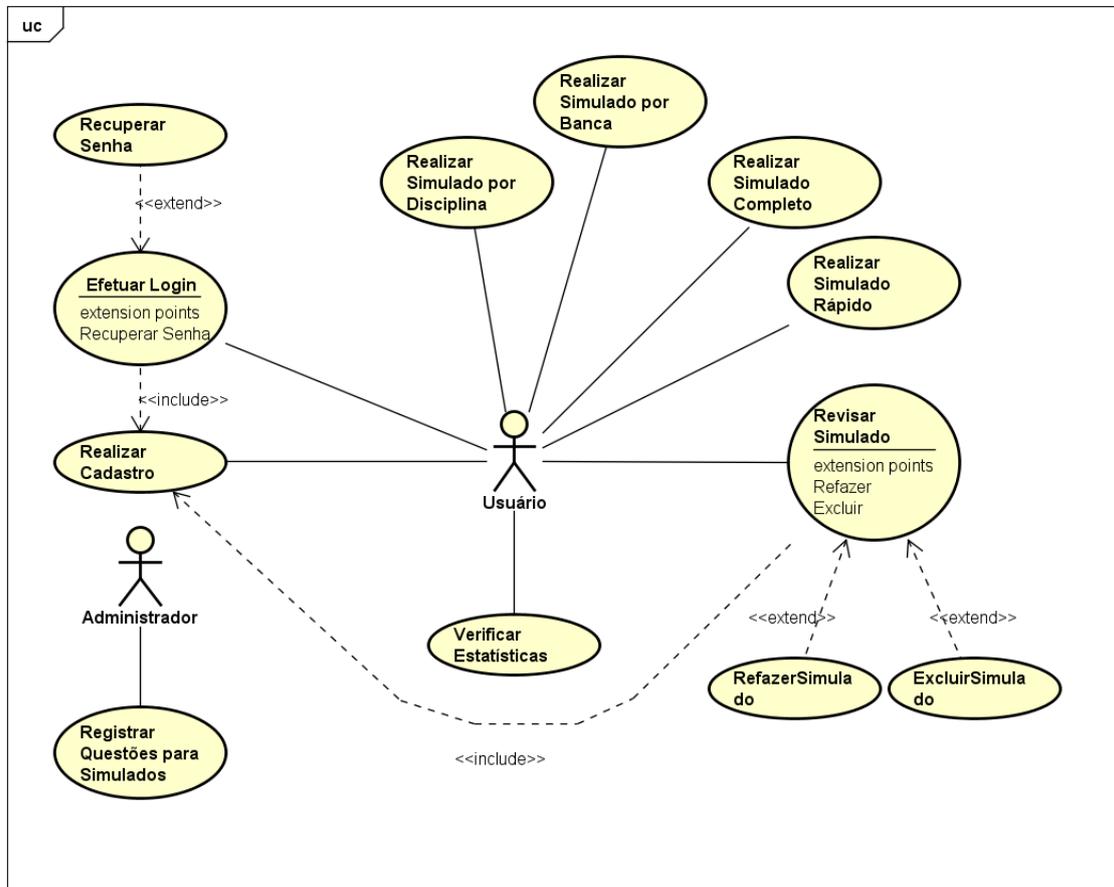


Figura 2. Diagrama de Casos de Uso.

O ator usuário pode realizar um cadastro e, se cadastrado com sucesso, pode efetuar um login no sistema. Ao ter seu login autenticado, o usuário pode realizar um simulado, verificar o seu desempenho geral nos simulados ou revisar um simulado já feito. Durante a revisão do simulado, é possível verificar o desempenho nele e refaze-lo, caso desejar.

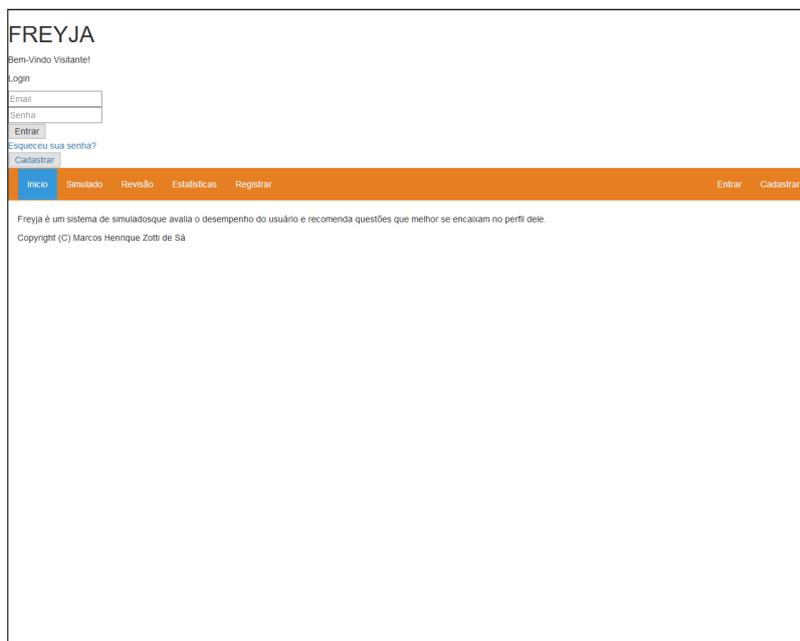


Figura 3. Tela de Início.

Na Figura 3 observamos a tela inicial do sistema proposto. Aqui é onde todos os usuários iniciam e escolhem as ações que desejam realizar. No topo da página é exibido o título do sistema e uma caixa de login com a opção de cadastro e recuperação de senha. Essa caixa persiste em todas as outras páginas, para que o usuário possa escolher fazer o login a qualquer instante. Logo abaixo encontra-se uma barra de navegação com a indexação de todas as outras páginas do sistema. Essa barra também está presente nas demais páginas. O corpo da página inicial apenas apresenta uma descrição do sistema.



Figura 4. Tela de Cadastro.

A Figura 4 é a tela de cadastro de usuário do sistema. Ela possui um simples formulário para ser preenchido. Todos os campos são obrigatórios.

FREYJA

Bem-Vindo Visitante!

Login

Email

Senha

Entrar

Esqueceu sua senha?

Cadastrar

Inicio Simulado Revisão Estatísticas Registrar

Escolha uma opção de Simulado para iniciar.

Simulado Completo

Iniciar

Simulado Rapido

Iniciar

Simulado por Disciplina

Matemática ▾ Iniciar

Copyright (C) Marcos Henrique Zotti de Sá

Figura 5. Tela de Iniciar Simulado.

A tela de início do simulado é demonstrada na Figura 5, nela temos as três opções de iniciar um simulado. Simulado completo – essa é a opção principal, o usuário inicia uma série de questões de diversas matérias. Simulado rápido – inicia-se um simulado com apenas 15 questões e sem a presença de um contador de tempo. Simulado por disciplina – esta opção consiste em um simulado que apresenta apenas questões da disciplina escolhida, dentre elas matemática, português, etc. E por fim, o Simulado por Banca, onde serão respondidas questões de determinadas bancas de concursos.

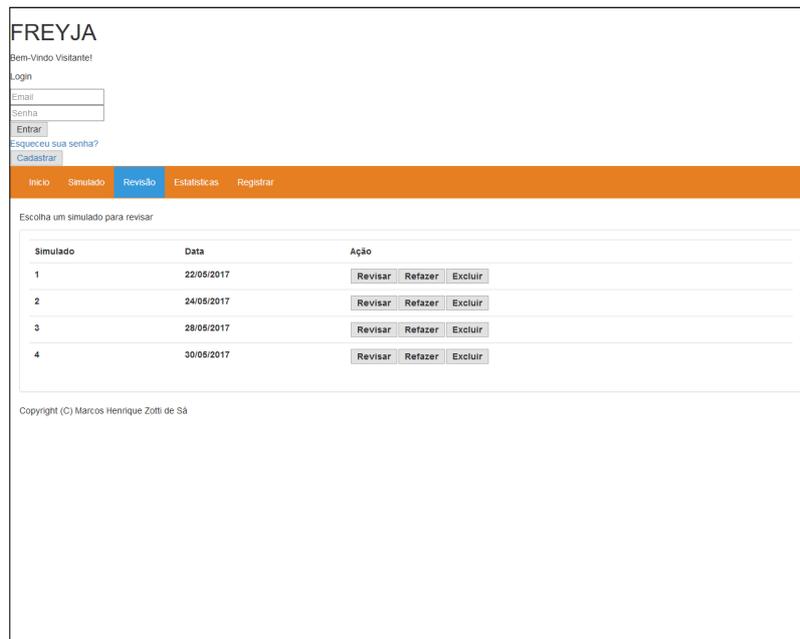


Figura 6. Tela de Revisão.

Para que o usuário possa revisar seus simulados anteriores, ele pode ir para a página demonstrada na Figura 6. O sistema exibe uma tabela com os simulados já realizados pelo usuário e a data em que eles foram respondidos. Ao lado existe uma coluna de ações, com as opções de: i) Revisar: exibe todas as questões do simulado com as respostas marcadas pelo usuário e as alternativas corretas; ii) Refazer: inicia-se o mesmo simulado uma outra vez com as mesmas questões; iii) Excluir: remove do sistema o simulado escolhido.

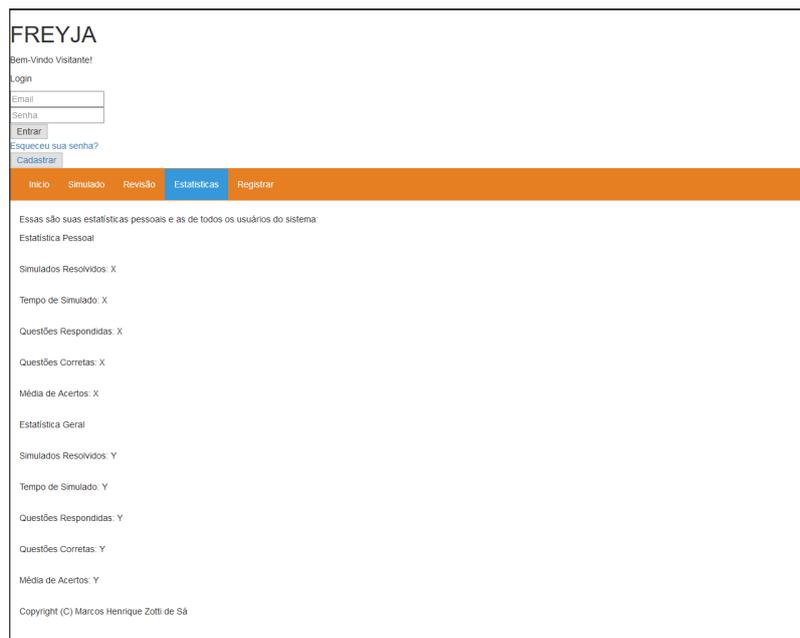


Figura 7. Tela de Estatísticas.

A seção de estatísticas ilustrada pela Figura 7 exibe dados de todos usuários do sistema e o usuário que estiver logado como simulados resolvidos, questões respondidas, média, etc.

FREYJA
Bem-Vindo Visitante!
Login
Email
Senha
Entrar
Esqueceu sua senha?
Cadastrar

Inicio Simulado Revisão Estatísticas Registrar

Cadastro de Questão
Preencha o formulário para realizar o cadastro de uma nova questão de concurso.

Titulo
Banca
Disciplina
Matéria
Dificuldade
Enunciado
Alternativa Correta
Alternativa Errada
Alternativa Errada
Alternativa Errada
Alternativa Errada
Cadastrar

Copyright (C) Marcos Henrique Zotti de Sá

Figura 8. Tela de Registro.

Para inserir novas questões no sistema, um usuário administrador deve preencher o formulário da Figura 8. O administrador deve preencher todos os campos com atenção para não ocorrer problemas na hora da resposta do usuário.

Com a grande quantidade de dados inseridos no sistema (usuários e questões), foi possível realizar a mineração de dados. Para isso foi utilizado o software RapidMiner, que realiza a mineração utilizando diversas operações. Foi utilizado o método de classificação onde os usuários foram divididos em 5 grupos, utilizando como critério a média deles, para que cada um se encaixe no grupo correto e recebam as questões melhores recomendadas quando iniciarem um simulado.

5. Resultados

Após o desenvolvimento do sistema, foram obtidos resultados relevantes para o encerramento do trabalho, apesar de dificuldades encontradas. Na etapa de desenvolvimento foram finalizadas as páginas web e o código fonte conforme foi citada na modelagem do trabalho. A funcionalidade principal está presente no sistema (realizar simulados), porém inicialmente seriam implementadas outras funcionalidades que serão citadas nos trabalhos futuros. Através da ferramenta RapidMiner, foi realizada a mineração de dados no banco Oracle, onde foi possível utilizar uma técnica de mineração em uma entrada de dados e manipular ela para melhor servir o sistema.

6. Metodologia

A metodologia utilizada no desenvolvimento deste trabalho foi:

1. Levantamento Bibliográfico: nesta etapa foram estudados os conceitos sobre sistemas de recomendação e mineração de dados.
2. Pesquisa Exploratória: nesta etapa foram pesquisados três sistemas com características semelhantes ao proposto. Essa etapa foi importante para definir as funcionalidades do sistema apresentado.
3. Definição das funcionalidades: nesta etapa foi definido como opera as funcionalidades do sistema com o sistema de recomendações, e a mineração de dados.
4. Estudo das tecnologias: nesta etapa foram realizadas pesquisas de quais linguagens de programação, frameworks e bancos de dados que seriam aplicados no desenvolvimento do trabalho. Após este estudo, foram escolhidas as seguintes tecnologias: i) Linguagem de programação java para o desenvolvimento do código fonte do sistema. Essa escolha foi feita devido a necessidade de uma linguagem com as características do paradigma orientado a objetos e a sua capacidade; ii) Banco de dados Oracle. Esta escolha se deu pelas características do banco e pela facilidade de poder interagir com a linguagem de programação java; iii) RapidMiner para realizar a mineração de dados.
5. Desenvolvimento do sistema: nesta etapa foi desenvolvido o sistema, iniciando pelo cadastro de usuários e as questões para serem usadas nos testes; foram desenvolvidas as funcionalidades tanto dos usuários sem cadastro quanto os que já possuem. O processo de desenvolvimento foi interativo e incremental, dessa forma, conforme as funcionalidades foram sendo desenvolvidas, as mesmas serão testadas.
6. Refinamento: após concluída a avaliação do sistema, foram realizadas ajustes no sistema.

7. Considerações Finais

Para este trabalho, foi realizado uma pesquisa dos sistemas de resolução de simulados utilizados no Brasil. O artigo apresentou as semelhanças entre esses sistemas e o proposto. Também foram pesquisados os conceitos de sistemas de recomendações, mineração de dados.

Os resultados obtidos foram relevantes e o sistema funciona como o proposto e esperado na modelagem, apesar de algumas dificuldades obtidas no desenvolvimento que ocasionaram na falta de alguma funcionalidades a mais que o sistema possuía.

A fim de trazer uma inovação para o tema de sistemas de resolução de simulados, o trabalho proposto apresenta uma solução que aplica os dados coletados no desempenho do usuário e faz recomendações com base nessa avaliação.

Para o aperfeiçoamento do sistema, em trabalhos futuros seriam executadas as seguintes implementações: Inserir as demais opções de simulados além do completo como o por matéria ou por banca, onde serão filtradas questões à desejo do usuário. É desejável também a implementação do sistema em plataformas mobile, que daria um maior alcance da

Referências

Cazella, S. C. (2011). A ciência da opinião: Estado da arte em sistemas de recomendação.

Coelho, K. R. (2015). Business intelligence (b.i.): Análise preditiva através da mineração de dados.

Fayyad, U. (1996). From data mining to knowledge discovery in databases.

Franco, M. H. I. (2016). Desenvolvimento de software iii.

Sferra, H. H. (2003). Conceitos e aplicações de data mining.