

Daniel Schreiber Guimarães

**Sistema de recomendação de
disciplinas para matrícula de um
aluno da PUC-Rio**

PROJETO FINAL

DEPARTAMENTO DE INFORMÁTICA
Programa de Graduação em Engenharia da
Computação

Rio de Janeiro
Novembro de 2023



Daniel Schreiber Guimarães

**Sistema de recomendação de disciplinas para
matrícula de um aluno da PUC-Rio**

Relatório de Projeto Final I

Relatório de Projeto Final, apresentado ao Programa de Engenharia da Computação, do Departamento de Informática da PUC-Rio como requisito parcial para a obtenção do título de Bacharel em Engenharia da Computação.

Orientador: Prof. Marcos Vianna Villas

Rio de Janeiro
Novembro de 2023

Todos os direitos reservados. A reprodução, total ou parcial do trabalho, é proibida sem a autorização da universidade, do autor e do orientador.

Daniel Schreiber Guimarães

Graduando em Engenharia da Computação na PUC - Rio

Ficha Catalográfica

Guimarães, Daniel Schreiber

Sistema de recomendação de disciplinas para matrícula de um aluno da PUC-Rio / Daniel Schreiber Guimarães; orientador: Marcos Vianna Villas. – 2023.

73 f: il. color. ; 30 cm

Projeto Final - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática, 2023.

Inclui bibliografia

1. Informática – Trabalho de Conclusão de Curso (Graduação). 2. Recomendação. 3. Disciplinas. I. Villas, Marcos Vianna. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Para os meus pais, que
nunca deixaram de me ajudar.

Abstract

Guimarães, Daniel Schreiber; Villas, Marcos Vianna (Advisor). **Sistema de recomendação de disciplinas para matrícula de um aluno da PUC-Rio**. Rio de Janeiro, 2023. 73p. Projeto Final – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

PUC-Rio allows students to choose which courses they may take in the next semester, but decision-making is not an easy task, given that relevant information such as the difficulty of the subject, available schedules, or available professor is not trivial to find. This graduation project is related to a system that provides a platform for students to build their schedules, evaluate courses and professors, and receive recommendations for courses to be taken.

Keywords

Recommendation; Courses.

Resumo

Guimarães, Daniel Schreiber; Villas, Marcos Vianna. **Sistema de recomendação de disciplinas para matrícula de um aluno da PUC-Rio**. Rio de Janeiro, 2023. 73p. Projeto Final – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

A PUC-Rio permite que os alunos escolham quais disciplinas irão cursar no próximo semestre, mas a tomada de decisão não é uma tarefa fácil, visto que as informações relevantes como dificuldade da matéria, horários disponíveis ou professores disponíveis não são triviais de se encontrar. Esse projeto de graduação é referente a um sistema que oferece uma plataforma para alunos montarem suas grades, avaliarem disciplinas e professores, e receberem recomendações de disciplinas para serem cursadas.

Palavras-chave

Recomendação; Disciplinas.

Sumário

| | | |
|-----------|---|-----------|
| 1 | Introdução | 11 |
| 2 | Situação Atual | 12 |
| 2.1 | Serviços disponibilizados | 12 |
| 2.2 | Processo de matrícula | 13 |
| 2.3 | Estudos relacionados | 14 |
| 3 | Objetivo | 15 |
| 4 | Plano de Ação | 16 |
| 4.1 | Cronograma original | 16 |
| 4.2 | Cronograma real | 16 |
| 5 | Validação do problema | 19 |
| 5.1 | Método das entrevistas | 19 |
| 5.2 | Perguntas realizadas | 19 |
| 5.3 | Resultados das entrevistas | 20 |
| 6 | Requisitos | 25 |
| 6.1 | Do algoritmo | 25 |
| 6.2 | Do sistema | 25 |
| 7 | Casos de Uso | 28 |
| 8 | Wireframe | 34 |
| 9 | Dados do sistema | 38 |
| 9.1 | Diagrama de entidades e relacionamentos | 38 |
| 9.2 | Modelo lógico | 39 |
| 9.3 | Dicionário de dados | 40 |
| 10 | Construção | 45 |
| 10.1 | Tecnologias utilizadas | 45 |
| 10.2 | Coleta de dados | 46 |
| 10.3 | Definição e implementação do algoritmo | 52 |
| 10.4 | Implementação da API | 57 |
| 10.5 | Implementação dos microserviços | 57 |
| 10.6 | Implementação da interface | 58 |
| 11 | Testes | 64 |
| 11.1 | Cobertura dos requisitos | 64 |
| 11.2 | Testes da interface | 66 |
| 11.3 | Teste de performance | 67 |
| 12 | Uso real | 68 |

| | |
|--------------------------------------|-----------|
| 13 Conclusão | 70 |
| 13.1 Estudos futuros | 70 |
| 13.2 Código completo | 71 |
| 14 Referências bibliográficas | 72 |

Lista de figuras

| | | |
|--------------|---|----|
| Figura 2.1 | Interface do <i>microhorario</i> | 12 |
| Figura 2.2 | Interface do simulador | 13 |
| Figura 4.1 | Fluxo das etapas realizadas | 17 |
| Figura 4.2 | Cronograma original do projeto | 18 |
| Figura 4.3 | Cronograma atualizado do projeto | 18 |
| Figura 5.1 | Distribuição da quantidade de entrevistados pelos seus períodos atuais. | 21 |
| Figura 7.1 | Diagrama dos casos de uso | 28 |
| Figura 8.1 | Wireframe da página inicial | 34 |
| Figura 8.2 | Wireframe da página de criação de grade horária | 35 |
| Figura 8.3 | Fluxograma da interface de criação de grade horária | 36 |
| Figura 8.4 | Wireframe da página de avaliação de disciplinas e professores | 37 |
| Figura 9.1 | Diagrama de entidades e relacionamentos | 39 |
| Figura 9.2 | Diagrama do modelo lógico dos dados | 40 |
| Figura 10.1 | Diagrama da arquitetura do sistema | 45 |
| Figura 10.2 | Seleção de uma disciplina com co-requisitos durante o processo de matrícula | 52 |
| Figura 10.3 | Documentação da API | 59 |
| Figura 10.4 | Implementação da tela inicial da interface | 60 |
| Figura 10.5 | Implementação da tela de autenticação do usuário | 60 |
| Figura 10.6 | Implementação da tela de cadastro de histórico | 61 |
| Figura 10.7 | Implementação da tela de criação de grade da interface | 61 |
| Figura 10.8 | Detalhes de uma disciplina na interface, exibido ao selecionar uma disciplina | 62 |
| Figura 10.9 | Implementação da tela de avaliação da interface | 62 |
| Figura 10.10 | Legenda da interface de criação de grade horária | 63 |
| Figura 11.1 | Teste de performance da recomendação | 67 |

Lista de tabelas

| | | |
|------------|--|----|
| Tabela 5.1 | Graus dos critérios de escolhas de disciplinas escolhidos pelos entrevistados | 22 |
| Tabela 5.2 | Relação de peso para grau dos critérios de escolhas | 23 |
| Tabela 5.3 | Pesos dos critérios de escolhas de disciplinas escolhidos pelos entrevistados. | 23 |
| Tabela 6.1 | Requisitos do algoritmo | 26 |
| Tabela 6.2 | Requisitos do sistema | 27 |
| Tabela 7.1 | Caso de uso UC01 | 29 |
| Tabela 7.2 | Caso de uso UC02 | 30 |
| Tabela 7.3 | Caso de uso UC03 | 30 |
| Tabela 7.4 | Caso de uso UC04 | 32 |
| Tabela 7.5 | Caso de uso UC05 | 33 |
| Tabela 9.1 | Dicionário de dados | 44 |

Lista de Códigos

| | | |
|----------|---|----|
| Código 1 | Modelo físico | 46 |
| Código 2 | Exemplo do arquivo de carga de currículos | 49 |
| Código 3 | Consulta dos dados para o algoritmo | 54 |

1

Introdução

A cada semestre, um aluno que está fazendo graduação no Departamento de Informática da PUC-Rio precisa fazer a sua matrícula em disciplinas oferecidas pela universidade para os próximos seis meses. Apesar de algumas restrições como pré-requisitos de disciplinas, o aluno tem ampla liberdade de escolher as disciplinas que melhor se encaixam na sua grade horária. Essa liberdade é uma vantagem devido à disponibilização de uma grade horária flexível para o aluno, porém precisa de um maior esforço de pesquisa e organização deste aluno para que a sua matrícula no próximo período seja feita de forma eficaz segundo critérios do próprio aluno.

A motivação desse projeto é o estudo, planejamento e desenvolvimento de um sistema de recomendação de disciplinas para o próximo período que auxilie o aluno em sua matrícula ao sugerir interativamente disciplinas com base em dados fornecidos pela universidade e por avaliações informadas por alunos.

2

Situação Atual

2.1

Serviços disponibilizados

O microhorario¹ é um serviço disponibilizado pela universidade para consultar informações das disciplinas do semestre atual. Este possui uma interface simples que pode ser visualizada na figura 2.1. Essa interface permite filtrar disciplinas por seus atributos como nome, código, professor e departamento. Porém, essa interface não é personalizada com relação ao aluno, ou seja, não exibe as disciplinas que o aluno em específico ainda não cursou ou que o aluno não pode cursar devido à pré-requisitos.

The screenshot displays the 'Horário das Disciplinas (sujeito a alteração) - 2023.1' page on the PUC-Rio website. The interface includes a search bar at the top right with the text 'Buscar em puc-rio.br' and a dropdown menu for 'Índice de A a Z'. Below the header, there are two orange banners: 'Aviso importante para alunos de Engenharia e Informática! Clique aqui.' and 'Atenção! Verifique aqui as disciplinas com oferecimento anual.' The main search form is titled 'Consulta' and contains several filters: 'Nível' (Sem Distinção), 'Código da Disciplina' (empty), 'Nome da Disciplina' (empty), 'Nº de Créditos' (empty), 'Nome do Professor' (empty), 'Destino (bloqueio)' (Qualquer destino), 'Dia da Semana' (Qualquer dia), 'Hora - Início Aula' (empty), 'Hora - Término Aula' (empty), and 'Disciplina sem horário fixo' (checkbox). There are also checkboxes for 'Disciplina com horas a distância' and 'Turno' (Qualquer turno). The 'Departamento da Disciplina/Turma' is set to 'Qualquer departamento'. At the bottom of the form are 'Buscar' and 'Limpar' buttons. The footer of the page provides contact information for PUC-Rio, including the address 'Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil', CEP '22451-900', and phone numbers '(55 21) 3527-1001 / 3736-1001'. It also mentions '31 anos na WEB' and 'Todos os direitos reservados'.

Figura 2.1: Interface do *microhorario*

Para consultar a grade recomendada e os pré-requisitos, o aluno de engenharia de computação precisa acessar um documento² presente na página do departamento ou acessar o currículo disponibilizado na página da universidade. Outros cursos disponibilizam documentos semelhantes em páginas diferentes. A grade recomendada possui os nomes das disciplinas e seus pré-requisitos, mas não contém a disponibilidade do próximo período.

Além disso, as disciplinas eletivas que o departamento oferece durante o semestre são anunciadas em diferentes veículos de comunicação, como e-mail, página da coordenação do curso, ou folhetos em corredores do departamento. Não há uma distribuição centralizada das ofertas de disciplinas eletivas,

¹<https://www.puc-rio.br/microhorario>

²http://www.inf.puc-rio.br/wordpress/wp-content/uploads/2023/01/Grade_Eng_comp_2023.pdf

portanto o aluno pode não saber aonde procurar essas ofertas, e perder boas oportunidades por desconhecer o anúncio da disciplina eletiva.

No final de cada período, a universidade oferece um serviço de avaliação de disciplinas e professores, permitindo avalia-los em várias categorias. Porém, os resultados das avaliações não são disponíveis publicamente.

2.2

Processo de matrícula

A matrícula na PUC-Rio é um processo que dura vários dias. A universidade divulga o agendamento de matrícula, em que cada aluno recebe uma data e hora para realizar sua matrícula através do portal online do aluno. As datas possíveis abrangem um período de cinco dias. Durante o período de matrícula, o microhorário atualiza periodicamente para indicar quais turmas ainda possuem vagas disponíveis, e se houve o cadastro ou cancelamento de alguma outra disciplina durante este período.

Portanto, é comum que um aluno crie formas de planejamento próprias para conseguir organizar todo o fluxo de dependências, anotar as disciplinas que estão sendo oferecidas, assim como suas turmas e professores. Esses dados precisam estar em constante atualização durante o período de matrícula, conforme as vagas vão sendo preenchidas e a disponibilização de novas turmas ou disciplinas são anunciadas. Este é um processo que maximiza a qualidade da sua grade horária, mas demanda muito esforço e tempo.

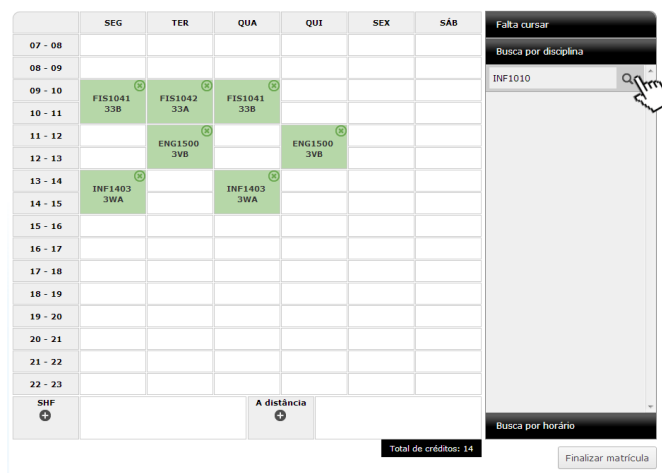


Figura 2.2: Interface do simulador

A universidade disponibiliza por dois ou três dias um simulador de matrícula antes do processo de matrícula em si. A interface do simulador é a mesma da matrícula, como pode ser observado na figura 2.2, afim de o aluno poder simular a criação da sua grade de disciplinas para o próximo semestre. O simu-

lador apresenta os dados das disciplinas atualizados conforme o microhorário, e os disponibiliza de três formas diferentes: buscas por disciplinas que faltam cursar no currículo do aluno, buscar por nomes de disciplinas e buscar por horários das turmas das disciplinas.[1]

2.3

Estudos relacionados

A escolha de disciplinas é um problema comum em universidades que possuem algum grau de flexibilidade na grade horária. Ng & Linn[2] desenvolveram um sistema de recomendação de disciplinas para a sua universidade chamado CrsRecs que utiliza análise de sentimento, pontuações de professores e disciplinas, e preferências que o aluno escolhe fornecer. A vantagem do CrsRecs é não depender de dados privados da universidade como notas e avaliação dos alunos nas disciplinas, utilizando somente informações fornecidas diretamente pelo usuário. Mas há também um esforço na decodificação das informações fornecidas, o que pode ocasionar recomendações incorretas provenientes da má interpretação do usuário.

Há estudos que utilizam históricos escolares dos alunos para gerar modelos de previsão de notas.[3, 4, 5] Nesse caso, o objetivo do modelo é tentar prever quais disciplinas o aluno tem a maior chance de obter uma boa nota, mas não necessariamente prever quais delas mais combina com o aluno.[3, 4]

A escolha de disciplinas eletivas também é uma preocupação presente nas universidades. Na PUC-Rio e em outras universidades, os currículos de graduação solicitam créditos de disciplinas eletivas, seja dentro do departamento da graduação do aluno ou não. O sistema do estudo de Adak et al.[5] recomenda disciplinas dentro de um departamento que parecem se relacionar com o histórico escolar do aluno, mas excluem disciplinas fora do departamento. Já o sistema de Xu et al.[6] recomenda uma sequência de disciplinas eletivas que maximem a sua nota e não aumentem o tempo de conclusão da graduação do aluno, levando em conta dados históricos do aluno e a disponibilidade atual das disciplinas. Por último, o estudo de Adak & Ercan[7] utiliza dois algoritmos de inteligência artificial de aprendizado supervisionado, support vector machine (SVM) e árvores de decisão, para recomendar eletivas que mais se combinam com o aluno, utilizando dados históricos do aluno tanto para o treinamento dos modelos de algoritmo como para a recomendação.

O comum nos estudos é a tentativa de maximizar a nota média do aluno utilizando os dados históricos de alunos da universidade para treinar modelos de inteligência artificial.[3, 4, 5, 7] Porém, o foco específico nos resultados pode ocultar oportunidades de disciplinas que o aluno poderia se interessar.

3

Objetivo

O sistema desenvolvido nesse trabalho consiste em uma interface de planejamento de matrícula semelhante ao simulador de matrícula integrado com um algoritmo de recomendação. O sistema disponibiliza as disciplinas conforme o microhorário, que durante o período exibe informações relativos ao período atual, e durante o final de um período e nas férias, exibe informações relativas ao próximo período. Para que haja uma personalização na recomendação das disciplinas, o sistema permite que o aluno carregue o histórico escolar na universidade afim de fornecer o histórico das disciplinas e seus graus para o algoritmo de recomendação.

O algoritmo de recomendação do sistema recebe as disciplinas oferecidas no próximo período e seus pré-requisitos, o modelo de grade recomendada pelo departamento e o histórico do aluno, e então responde com disciplinas selecionadas. O algoritmo também pode ser personalizado com informações de preferência do usuário.

O sistema armazena o histórico fornecido pelo aluno sem seus dados pessoais, afim de montar uma base de dados contendo exemplos de grades de alunos e suas notas, para que o algoritmo a utilize para obter resultados mais satisfatórios para o aluno usuário do sistema.

O sistema e o algoritmo foi desenvolvido especificamente para alunos do Departamento de Informática da PUC-Rio devido a dificuldade de abranger todas os currículos nos diferentes departamentos da universidade. Essa restrição também permitiu restringir o escopo do projeto afim de tentar obter melhores resultados. Os cursos considerados são Engenharia de Computação e Ciência da Computação.

4

Plano de Ação

Antes de especificar o sistema e o algoritmo, foi realizada uma validação do problema. Alunos do Departamento de Informática da universidade foram entrevistados afim de descobrir quais dificuldades enfrentam durante o processo de matrícula e quais são as suas preferências ao montar uma grade disciplinar. O objetivo era identificar quais características de uma grade disciplinar contribuem para a satisfação do aluno e adicionar as funcionalidades necessárias no sistema e no algoritmo.

Com o problema validado, então o projeto passou por uma etapa de estudo anteriores. Foram estudados diferentes algoritmos de recomendação em contextos semelhantes ao sistema a ser desenvolvido, afim de se projetar o algoritmo que mais se adequa às necessidades do problema validado. Nessa etapa, também foram analisadas as fontes de informações disponibilizadas pela universidade e como cada fonte pode ser integrada no sistema.

Após a etapa de estudos, deu-se início ao processo de modelagem dos requisitos, que foi seguido das elaborações de casos de usos, da arquitetura dos dados e da interface.

Finalizada as elaborações e diagramas, o desenvolvimento foi iniciado, dando prioridade para o algoritmo. Depois de obter um algoritmo minimamente viável, foi desenvolvida a interface de planejamento para mostrar o funcionamento do algoritmo. Tanto o algoritmo com a interface foram desenvolvidas utilizando um modelo de desenvolvimento incremental, para que o sistema passasse por etapas de desenvolvimento, teste e validação com os alunos.

A figura 4.1 exemplifica o fluxo das etapas efetuadas, desde os estudos preliminares até os testes e trabalhos futuros.

4.1

Cronograma original

O cronograma inicial pode ser visualizado na figura 4.2.

4.2

Cronograma real

O cronograma real pode ser visualizado na figura 4.3. Surgiu uma etapa de elaboração de especificações, que dizem respeito às elaborações de diagramas e tabelas de casos de uso, arquitetura de dados e interface, criadas antes da

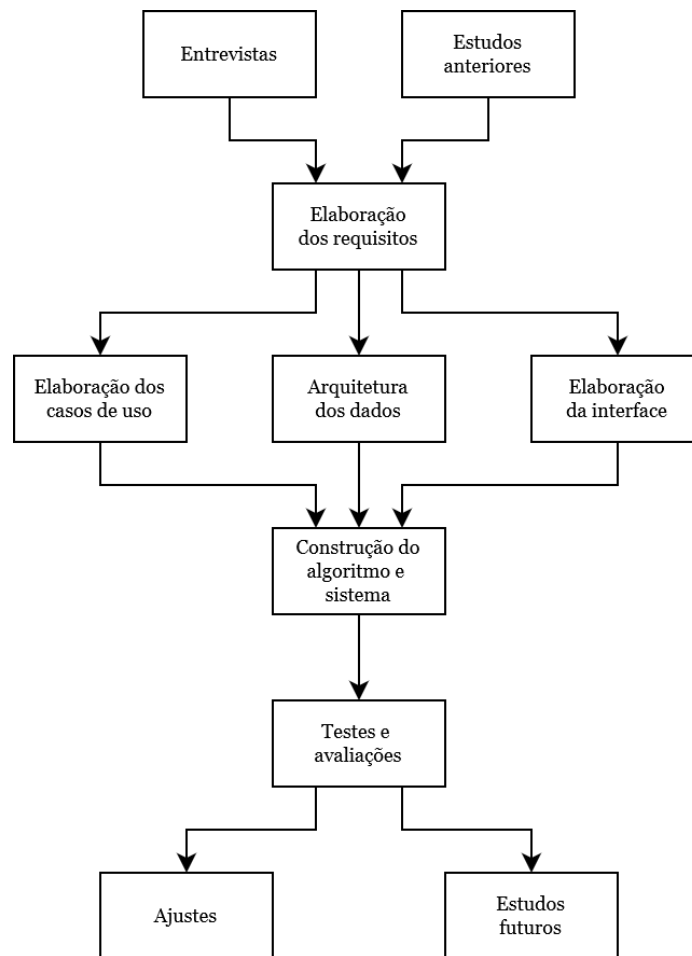


Figura 4.1: Fluxo das etapas realizadas

etapa de desenvolvimento. O desenvolvimento teve seu tempo encurtado de forma extrema, pois elaborar as especificações tomou grande parte do tempo de desenvolvimento, mas mesmo com metade do tempo original foi possível implementar todas as funcionalidades planejadas.

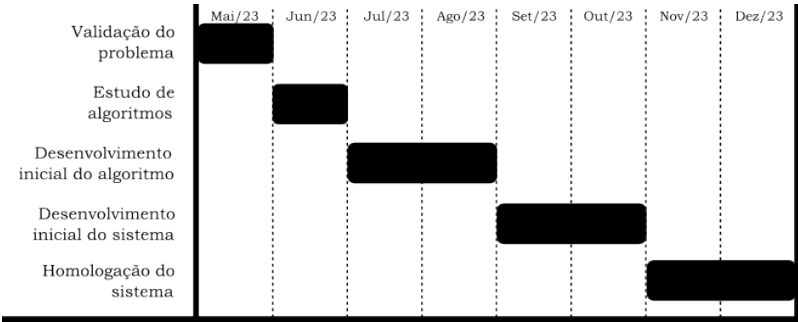


Figura 4.2: Cronograma original do projeto

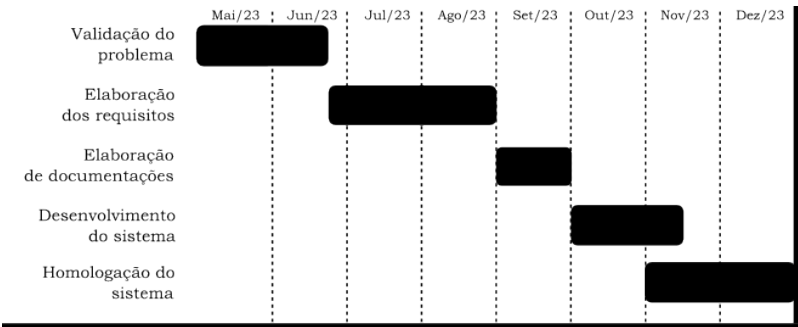


Figura 4.3: Cronograma atualizado do projeto

5

Validação do problema

A validação do problema foi efetuada por meio de entrevistas com alunos de graduação do Departamento de Informática. As entrevistas tinham como objetivo obter as opiniões de diferentes alunos em diferentes períodos de suas graduações quanto às suas preferências ao realizar as suas matrículas. As opiniões foram essenciais para planejar um algoritmo de recomendação que abrange as preferências do máximo de estudantes.

5.1

Método das entrevistas

Os entrevistados são alunos de graduação dos cursos do Departamento de Informática da PUC-Rio, cursando os cursos de engenharia da computação ou de ciência da computação. Foram selecionados vários alunos em diferentes estágios da graduação, do segundo ao sexto ano de graduação. Como algumas perguntas se referem às escolhas de disciplinas eletivas, e essas só começam a ser escolhidas no meio do período de graduação, não foram escolhidos estudantes no primeiro ano de graduação.

5.2

Perguntas realizadas

A seguir estão as seis perguntas realizadas para cada entrevistado do Departamento de Informática, assim como uma explicação do motivo da pergunta ser realizada.

1. **Qual é o seu curso e período atual?** O período atual do entrevistado é utilizado para entender o contexto de suas respostas, pois é necessário observar se o período em que o graduando está altera sua opinião ao realizar a sua matrícula. Como a quantidade de períodos propostos para a conclusão do curso variam de curso para curso no Departamento de Informática, também foi solicitado ao entrevistado seu curso para entender seu progresso de estudo na universidade e entender ainda melhor o contexto das suas próximas respostas.
2. **Em média quantos créditos você faz num semestre?** A quantidade de créditos por semestre é relevante para que o algoritmo de recomendação tenha um conhecimento da média de disciplinas selecionadas por período.

3. **Você se prepara com antecedência para sua matrícula? Se sim, com quanta antecedência? Se não, por quê?** Essa pergunta é fundamental para planejar uma expectativa de utilização da ferramenta de recomendação de disciplinas.
4. **Que critérios você utiliza para escolher uma disciplina? Dados os critérios fornecidos, coloque-os em uma ordem de mais relevante para menos relevante.** Essa é a pergunta mais fundamental para o planejamento do algoritmo. Afinal, o objetivo do algoritmo de recomendação é recomendar disciplinas relevantes para um aluno. Por isso, é necessário saber o que torna uma disciplina relevante para ser escolhida.
5. **Como você procura disciplinas eletivas?** O objetivo dessa pergunta é descobrir como cada estudante procura as disciplinas eletivas, para saber quais são as principais fontes de informações utilizadas.
6. **Você, pessoalmente, prefere disciplinas mais fáceis, mas que talvez não sejam muito úteis para a sua carreira, ou disciplinas mais relevantes mas que possam ser mais difíceis?** Essa pergunta pretende entender qual das duas opções é mais frequente. Caso a maioria dos estudantes preferem disciplinas mais fáceis, o algoritmo irá depender mais das notas dos alunos, e não tanto do conteúdo da disciplina em si. Caso a maioria dos estudantes preferem disciplinas mais relevantes, a nota pode não ser um valor tão relevante.

Ao final da entrevista, também foi oferecida a oportunidade de fornecer alguma opinião sobre o sistema de recomendação a ser desenvolvido.

5.3

Resultados das entrevistas

As seis perguntas foram separadas em três grupos, cada grupo com duas perguntas: *Dados básicos dos estudantes* (perguntas 1 e 2), *Escolha de disciplinas* (perguntas 3 e 4) e *Disciplinas eletivas* (perguntas 5 e 6). A seguir estão os resultados das perguntas, separados pelos respectivos grupos.

5.3.1

Dados básicos dos estudantes (Perguntas 1 e 2)

Dos onze entrevistados, cinco eram alunos de ciência de computação e os outros seis eram de engenharia de computação. cinco dos seis alunos entrevistados de engenharia de computação já haviam cursado mais da metade

do curso (já cursaram 5 dos 10 semestres recomendados), e três dos cinco alunos entrevistados de ciência de computação já haviam cursado mais da metade do curso (já cursaram 4 dos 8 semestres recomendados). Além disso, dos onze alunos, nove eram homens e dois eram mulheres.

A distribuição dos períodos atuais dos entrevistados pode ser visualizada na figura 5.1.

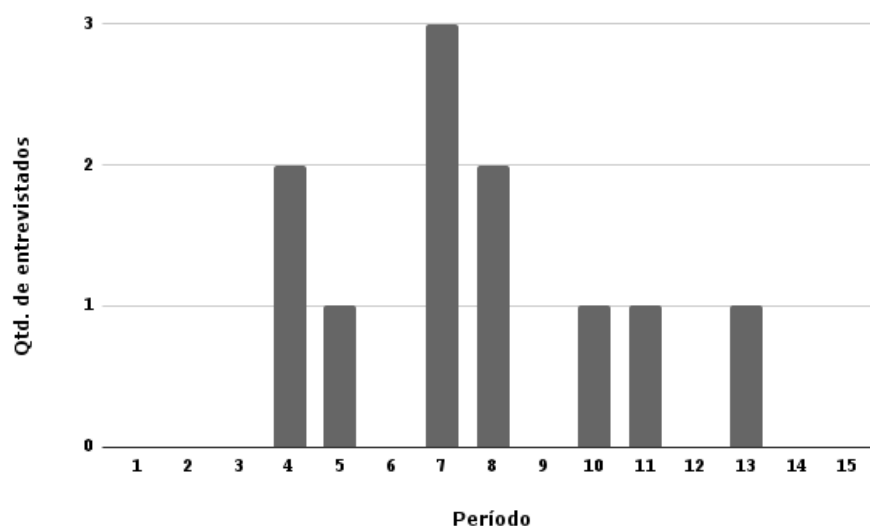


Figura 5.1: Distribuição da quantidade de entrevistados pelos seus períodos atuais.

A quantidade média de créditos por semestre dos alunos entrevistados é de 24 créditos, semelhante à quantidade média de créditos por semestre no currículo recomendado de ambos os cursos.

5.3.2

Escolha de disciplinas (Perguntas 3 e 4)

Dez dos onze entrevistados se preparam com antecedência para a matrícula. Três desses utilizam como base o repositório de disciplinas microhorário para buscar as informações das disciplinas oferecidas para o próximo período, e também utilizam o microhorário para descobrir quais são as disciplinas eletivas disponibilizadas no período. Todos os dez entrevistados utilizam o simulador de matrícula para confirmar suas escolhas. Apenas um dos entrevistados não se prepara com antecedência, fazendo suas pesquisas e escolhas durante o período da própria matrícula.

Os critérios preferidos de cada aluno variam bastante. Por exemplo, um entrevistado citou que alguns critérios são *"a matéria em si, se ela é mais difícil de entender ou mais fácil, e também o método de avaliação, ou seja,*

se a nota final é composta de somente duas provas, ou se são vários trabalhos pequenos espalhados pelo semestre. Além disso, eu diria que o horário também é um limitador, porque não consigo ter aula sete da manhã todos os dias. Mas eu acho que o mais importante para mim é o professor: Se eu sei quem é o professor, sei que ele explica bem, vale a pena mesmo que a matéria seja mais difícil. Eu diria que a ordem dos critérios seria o professor, depois a matéria em si, depois o método de avaliação e por último o horário da disciplina."

Os critérios de escolha de disciplinas ditos nas entrevistas foram agrupados em cinco grupos: (1) Conteúdo da disciplina; (2) Professor; (3) Método de avaliação; (4) Horário da disciplina e (5) Opinião de amigos. A tabela 5.1 apresenta as respostas dos onze entrevistados e as suas respostas. O número indica em qual posição ficou aquele critério na ordem de preferência pessoal do aluno.

| | Conteúdo | Professor | Avaliação | Horário | Opinião |
|------------------------|----------|-----------|-----------|---------|---------|
| <i>Entrevistado 1</i> | 3 | - | - | 1 | 2 |
| <i>Entrevistado 2</i> | 1 | - | - | 3 | 2 |
| <i>Entrevistado 3</i> | 2 | 3 | - | - | 1 |
| <i>Entrevistado 4</i> | 3 | 1 | - | 2 | - |
| <i>Entrevistado 5</i> | 2 | - | - | 1 | 3 |
| <i>Entrevistado 6</i> | 1 | - | - | - | - |
| <i>Entrevistado 7</i> | 1 | - | - | - | - |
| <i>Entrevistado 8</i> | 2 | 3 | 1 | 4 | - |
| <i>Entrevistado 9</i> | 3 | 2 | - | - | 1 |
| <i>Entrevistado 10</i> | 1 | - | 3 | 2 | - |
| <i>Entrevistado 11</i> | - | 2 | - | 1 | - |

Tabela 5.1: Graus dos critérios de escolhas de disciplinas escolhidos pelos entrevistados

Para gerar uma comparação dos critérios, foi dado um peso para cada grau de importância. Quanto maior o grau de importância. A tabela 5.2 exibe a relação entre os graus de importância e os pesos.

Ao substituir os pesos da Tabela 5.2 na Tabela 5.1, é possível somar os pesos para cada critério de escolha e obter uma relação entre eles, conforme a tabela 5.3.

Observando a Tabela 5.3, as somas dos pesos indicam que a ordem dos critérios mais relevantes na escolha de uma disciplina, do mais relevante para

| | | | | | |
|-------------|---|---|---|---|---|
| <i>Grau</i> | 1 | 2 | 3 | 4 | - |
| <i>Peso</i> | 5 | 4 | 3 | 2 | 0 |

Tabela 5.2: Relação de peso para grau dos critérios de escolhas

| | Conteúdo | Professor | Avaliação | Horário | Opinião |
|------------------------|----------|-----------|-----------|---------|---------|
| <i>Entrevistado 1</i> | 3 | 0 | 0 | 5 | 4 |
| <i>Entrevistado 2</i> | 5 | 0 | 0 | 3 | 4 |
| <i>Entrevistado 3</i> | 4 | 3 | 0 | 0 | 5 |
| <i>Entrevistado 4</i> | 3 | 5 | 0 | 4 | 0 |
| <i>Entrevistado 5</i> | 4 | 0 | 0 | 5 | 3 |
| <i>Entrevistado 6</i> | 5 | 0 | 0 | 0 | 0 |
| <i>Entrevistado 7</i> | 5 | 0 | 0 | 0 | 0 |
| <i>Entrevistado 8</i> | 4 | 3 | 5 | 2 | 0 |
| <i>Entrevistado 9</i> | 3 | 4 | 0 | 0 | 5 |
| <i>Entrevistado 10</i> | 5 | 0 | 3 | 4 | 0 |
| <i>Entrevistado 11</i> | 0 | 4 | 0 | 5 | 0 |
| Soma | 41 | 19 | 8 | 28 | 21 |

Tabela 5.3: Pesos dos critérios de escolhas de disciplinas escolhidos pelos entrevistados.

o menos relevante: *Conteúdo da disciplina*, *Horário da disciplina*, *Opinião de amigos*, *Professor* e *Método de avaliação*.

5.3.3

Disciplinas eletivas (Perguntas 5 e 6)

Dos onze entrevistados, nove citaram a utilização do microhorário para coletar as disciplinas eletivas do período. Os outros dois dependem mais de recomendação de amigos. Três entrevistados procuram por uma listagem oficial de disciplinas eletivas oferecidas pelo departamento para o próximo período, mas nem sempre encontram essa listagem oficial.

Por último, quatro entrevistados preferem disciplinas eletivas mais fáceis, enquanto três entrevistados preferem disciplinas eletivas que sejam mais relevantes para a sua formação. Os quatro restantes preferem um equilíbrio entre facilidade e relevância. Dois entrevistados disseram que inicialmente es-

colhem disciplinas mais relevantes, mas que ao decorrer dos períodos optam por eletivas mais fáceis. Uma das respostas foi que *"[Eu] prefiro [disciplinas] eletivas mais fáceis, porque eu as vejo como uma oportunidade para facilitar a minha vida na faculdade, pois já tenho outras disciplinas mais difíceis para me preocupar. Eu encaro as [disciplinas] eletivas como um escape para poder respirar."*

6

Requisitos

Os requisitos estão divididos em requisitos relacionados ao sistema e requisitos relacionados ao algoritmo. Os requisitos do sistema dizem respeito ao sistema que disponibiliza o serviço de montagem de grade horária do próximo período, assim como um serviço de avaliar disciplinas e professores, que existe para satisfazer as necessidades observadas nos capítulos anteriores. Os requisitos do algoritmo dizem respeito às entradas e saídas do algoritmo afim de se produzir uma recomendação adequada de acordo as necessidades observadas no capítulo 5.

6.1

Do algoritmo

Os requisitos do algoritmo de recomendação de disciplinas precisam satisfazer os critérios de escolha de disciplinas mais relevantes, discutidos no capítulo 5.3.2, ou seja, recomendar de acordo com o conteúdo da disciplina, horários disponíveis, opiniões de amigos, o professor, e por último o método de avaliação. Além disso, para satisfazer a preferência de disciplinas eletivas discutidas no capítulo 5.3.3, os requisitos precisam satisfazer a preferência de recomendar disciplinas com base na sua facilidade.

A lista de requisitos do algoritmo está disponível na tabela 6.1.

6.2

Do sistema

O sistema precisa satisfazer as necessidades do usuário e também as necessidades do algoritmo, pois o sistema hospeda o algoritmo de recomendação. Os requisitos do sistema estão disponíveis na tabela 6.2.

| | |
|-------------|--|
| RF01 | O algoritmo deve receber as disciplinas e turmas oferecidas no próximo semestre conforme o microhorário. |
| RF02 | O algoritmo deve receber opcionalmente a grade curricular do curso que o aluno está cursando. |
| RF03 | O algoritmo deve receber as avaliações das disciplinas e professores fornecidas pelos usuários do sistema. |
| RF04 | O algoritmo deve receber opcionalmente o histórico escolar do aluno. |
| RF05 | O algoritmo deve receber disciplinas já selecionadas pelo aluno. |
| RF06 | O algoritmo deve retornar recomendações de disciplinas com base nas entradas fornecidas. |
| RNF1 | O algoritmo deve retornar as recomendações em menos de 3 segundos. |
| RNF2 | O algoritmo deve ser determinístico, ou seja, retorna as mesmas recomendações para as mesmas entradas. |

Tabela 6.1: Requisitos do algoritmo

| | |
|-------------|--|
| RF07 | O sistema deve permitir que o usuário crie uma grade horária para o próximo período. |
| RF08 | O sistema deve permitir que o usuário submeta seu histórico escolar. |
| RF09 | O sistema deve permitir que o usuário selecione turmas das disciplinas para compor sua grade horária. |
| RF10 | O sistema deve permitir que o usuário autenticado armazene a sua grade horária finalizada. |
| RF11 | O sistema deve permitir que o usuário compartilhe a sua grade horária finalizada. |
| RF12 | O sistema deve permitir que o usuário recupere uma grade horária montada a partir de um link compartilhado. |
| RF13 | O sistema deve permitir que o usuário inicie uma sessão autenticada via o Sistema Acadêmico Universitário (SAU). |
| RF14 | O sistema deve permitir que o usuário finalize uma sessão autenticada. |
| RF15 | O sistema deve permitir que o usuário autenticado avalie uma disciplina. |
| RF16 | O sistema deve permitir que o usuário autenticado avalie um professor de uma disciplina. |
| RF17 | O sistema deve permitir que o usuário autenticado altere uma avaliação feita anteriormente. |
| RF18 | O sistema deve periodicamente sincronizar as vagas e turmas com o microhorário. |
| RF19 | O sistema deve periodicamente adicionar novas disciplinas conforme disponibilizado no microhorário. |
| RNF3 | O sistema deve suportar pelo menos 40 usuários simultâneos. |

Tabela 6.2: Requisitos do sistema

7

Casos de Uso

Os casos de uso do sistema servem para descrever como o sistema pode ser utilizado afim de se satisfazer as necessidades do usuário. Os casos de uso referenciam os requisitos do sistema descritos no capítulo 6. A figura 7.1 contém o diagrama de casos de uso do sistema.

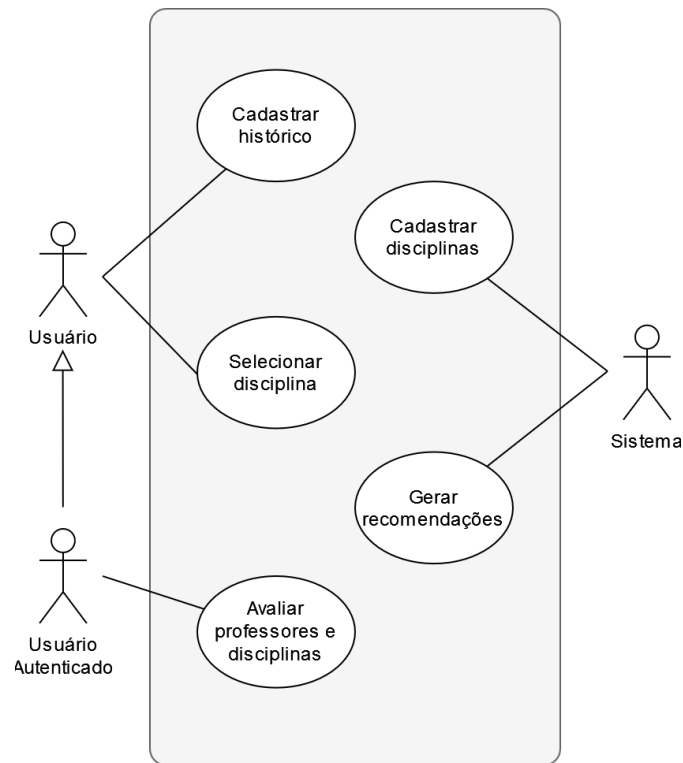


Figura 7.1: Diagrama dos casos de uso

A seguir estão as descrições dos casos de uso presentes na figura 7.1.

| Caso de Uso UC01 - Cadastrar disciplinas | |
|---|---|
| Objetivo | Permitir que o sistema atualize as disciplinas e suas informações no banco de dados, incluindo as eletivas. |
| Requisitos | RF01, RF09, RF18 e RF19 |
| Atores | Sistema |
| Pré condições | Não se aplica |

| Caso de Uso UC01 (continuação) | |
|---------------------------------------|--|
| Fluxo principal | [1] O sistema acessa o microhorário, coletando informações de todas as disciplinas. |
| | [2] O sistema trata e converte as informações para o modelo do banco. |
| | [3] O sistema armazena as novas informações no banco. |
| | [4] O sistema atualiza a data e hora da última atualização do microhorário que aparece na interface. |
| | [5] O caso de uso é encerrado. |

Tabela 7.1: Caso de uso UC01

| Caso de Uso UC02 - Cadastrar histórico | |
|---|--|
| Objetivo | Permitir que o usuário cadastre seu histórico escolar no sistema para personalizar as recomendações. |
| Requisitos | RF02, RF04, RF08 |
| Atores | Usuário |
| Pré condições | O usuário seleciona a opção de cadastrar histórico. |
| Fluxo principal | [1] O sistema exibe uma tela sobreposta solicitando o histórico escolar do aluno. |
| | [2] O usuário submete o histórico escolar. [A1] |
| | [3] O sistema armazena o histórico e o currículo do aluno. |
| | [4] O sistema exibe um texto de histórico cadastrado com sucesso. [A2] |
| | [5] O caso de uso é encerrado. |
| Fluxos Alternativos | [A1] O usuário pressiona fora da tela de cadastro de histórico |
| | [1] O caso de uso é encerrado. |

| Caso de Uso UC02 (continuação) | |
|---------------------------------------|--|
| | [A2] O sistema não consegue encontrar o currículo ou as disciplinas cursadas no currículo |
| | [1] O sistema exibe um texto de erro. |
| | [2] O sistema volta para o passo 1 do fluxo principal. |

Tabela 7.2: Caso de uso UC02

| Caso de Uso UC03 - Gerar recomendações | |
|---|---|
| Objetivo | Permitir que o sistema recomende disciplinas para o usuário. |
| Requisitos | RF01-06, RNF01-02 |
| Atores | Sistema |
| Pré condições | O usuário modificou a grade horária. |
| Fluxo principal | [1] O sistema coleta as disciplinas e turmas já adicionadas no grade horária do usuário. |
| | [2] O sistema recupera o histórico escolar do usuário do banco de dados, caso o usuário esteja autenticado e o tenha submetido.[A1] |
| | [3] O sistema utiliza o algoritmo para gerar recomendações de disciplinas para o usuário. |
| | [4] O sistema exibe as recomendações na interface do usuário |
| | [5] O caso de uso é encerrado. |

Tabela 7.3: Caso de uso UC03

| Caso de Uso UC04 - Selecionar disciplina | |
|---|--|
| Objetivo | Permitir que o usuário adicione e remova disciplinas da sua grade. |
| Requisitos | RF01, RF09-RF12 |
| Atores | Usuário |
| Pré condições | O usuário está na área de criação da grade horária. |
| Fluxo principal | [1] O sistema exibe a grade horária do usuário, uma lista de disciplinas disponíveis, uma lista de disciplinas recomendadas e um campo de texto para pesquisa. |
| | [2] O usuário seleciona uma disciplina da lista de disciplinas disponíveis ou recomendadas. [A1] [A2] . |
| | [3] O sistema exibe as turmas disponíveis para a disciplina selecionada e um botão de voltar. |
| | [4] O usuário seleciona uma das turmas exibidas. [A3] |
| | [5] O sistema acrescenta a turma selecionada na grade horária, e recalcula as disciplinas recomendadas. |
| | [6] O caso de uso é encerrado. |
| Fluxos Alternativos | [A1] O sistema seleciona uma das disciplinas na sua grade horária |
| | [1] O sistema exibe um botão de informação e um botão de excluir. |
| | [2] O usuário seleciona o botão de excluir. [A3] |
| | [3] O sistema remove a turma e disciplina da grade horária do usuário, e recalcula as disciplinas recomendadas. |
| | [4] O sistema volta para o passo 1 do fluxo principal. |
| | [A2] O usuário preenche o campo de texto para pesquisa |

| Caso de Uso UC04 (continuação) | |
|---------------------------------------|--|
| | [1] O sistema altera a lista de disciplinas, filtrando de acordo com o texto do campo de pesquisa. |
| | [2] O sistema volta para o passo 1 do fluxo principal. |
| | [A3] O usuário pressiona o botão de informação |
| | [1] O sistema volta para o passo 3 do fluxo principal. |

Tabela 7.4: Caso de uso UC04

| Caso de Uso UC05 - Avaliar disciplinas e professores | |
|---|--|
| Objetivo | Permitir que o usuário avalie uma disciplina ou um professor. |
| Requisitos | RF03, RF15 e RF16 |
| Atores | Usuário |
| Pré condições | Não se aplica. |
| Fluxo principal | [1] O sistema exibe uma lista de disciplinas e professores e um campo de texto para pesquisa. |
| | [2] O usuário seleciona uma disciplina. [A1] [A2] |
| | [3] O sistema exibe uma tela para avaliar a disciplina, um botão de salvar e um botão de voltar. |
| | [4] O usuário avalia a disciplina e seleciona o botão de salvar. [A3] |
| | [5] O sistema armazena a avaliação do usuário. |
| | [6] O caso de uso é encerrado. |
| Fluxos Alternativos | [A1] O usuário seleciona um professor |
| | [1] O sistema exibe uma tela para avaliar o professor, um botão de salvar e um botão de voltar. |

| Caso de Uso UC05 (continuação) | |
|---------------------------------------|--|
| | [2] O usuário avalia o professor e seleciona o botão de salvar. [A3] |
| | [3] O sistema armazena a avaliação do usuário. |
| | [4] O caso de uso é encerrado. |
| | [A2] O usuário preenche o campo de texto para pesquisa |
| | [1] O sistema exibe disciplinas e professores utilizando como filtro o texto do usuário. |
| | [2] O sistema volta para o passo 1 do fluxo principal. |
| | [A2] O usuário seleciona o botão de voltar |
| | [2] O sistema volta para o passo 1 do fluxo principal. |

Tabela 7.5: Caso de uso UC05

8 Wireframe

Um *wireframe* é um protótipo da interface do sistema, que serve para ilustrar o funcionamento e interação do usuário com o sistema. A seguir são apresentadas três telas do sistema: a página inicial, a página de criação de grade horária, e a página de avaliação de disciplinas e professores. Essas são as principais páginas do sistema.

O design da página inicial (*Landing Page*), apresentado na figura 8.1, busca ser simples, com uma rápida descrição do sistema e das suas funcionalidades, e com dois botões para enviar o usuário para a página de criação de grade ou de avaliação.

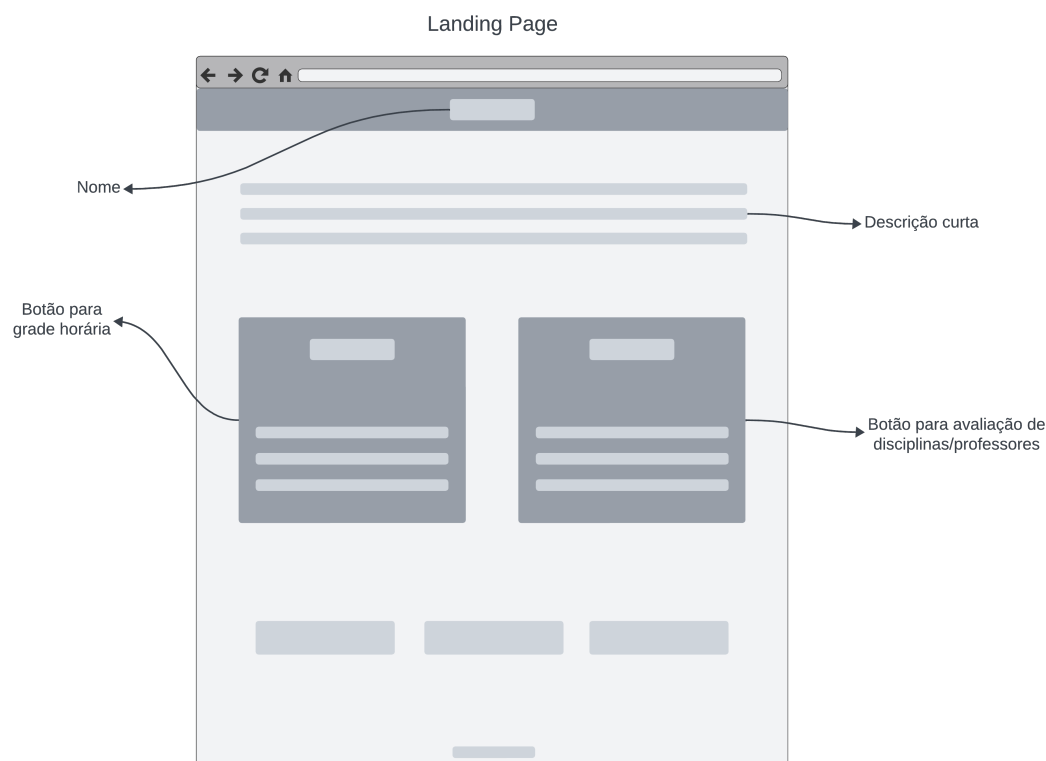


Figura 8.1: Wireframe da página inicial

A página de criação de grade horária, apresentado na figura 8.2, pretende ser semelhante à interface de matrícula real da PUC-Rio, como na figura 2.2, com a adição de mais informações que auxiliem o usuário a fazer suas escolhas, além da barra de disciplinas recomendadas.

Para melhor descrever a interação do usuário com o sistema com a interface da criação de grade horária e com o algoritmo de recomendação, a figura 8.3 apresenta um fluxograma com as possíveis ações do usuário na interface.

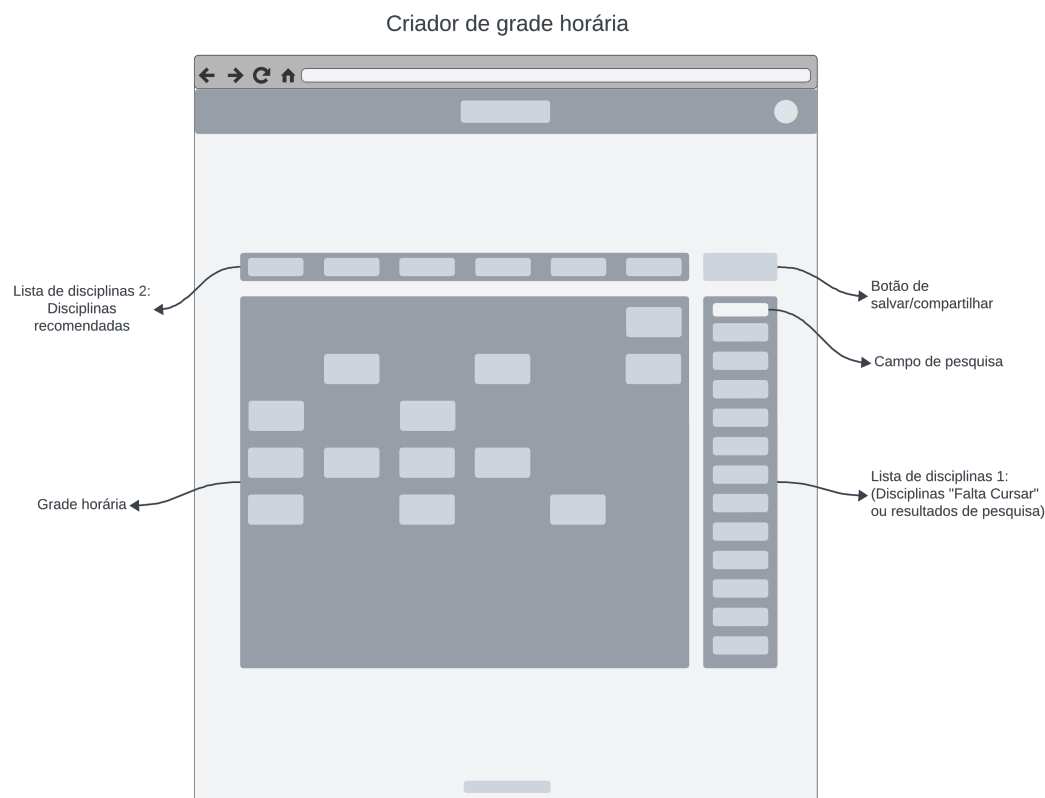


Figura 8.2: Wireframe da página de criação de grade horária

Neste diagrama, os retângulos representam componentes da interface passíveis de interação do usuário, os hexágonos representam operações efetuadas pelo sistema, e o losango representa uma tomada de decisão pelo sistema. As setas contínuas representam operações do usuário, e as setas tracejadas representam transições automáticas entre os componentes e operações.

Por fim, a página de avaliação de disciplinas e professores, apresentado na figura 8.4, procura ser uma interface simples, com uma lista das disciplinas cursadas ou professores, e com a opção de avaliar no formato de estrelas, sendo a pior avaliação uma estrela, e cinco estrelas a melhor.

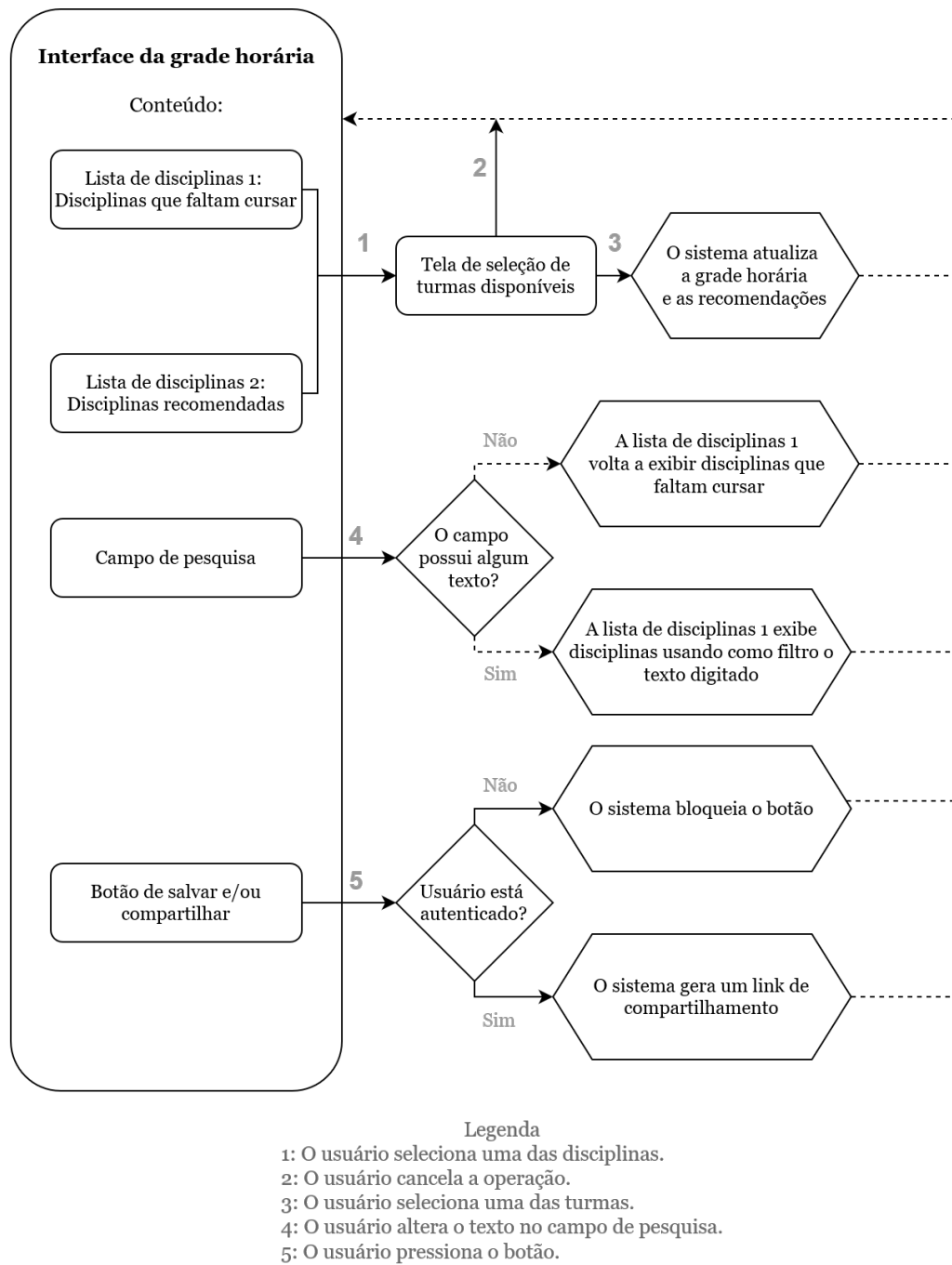


Figura 8.3: Fluxograma da interface de criação de grade horária

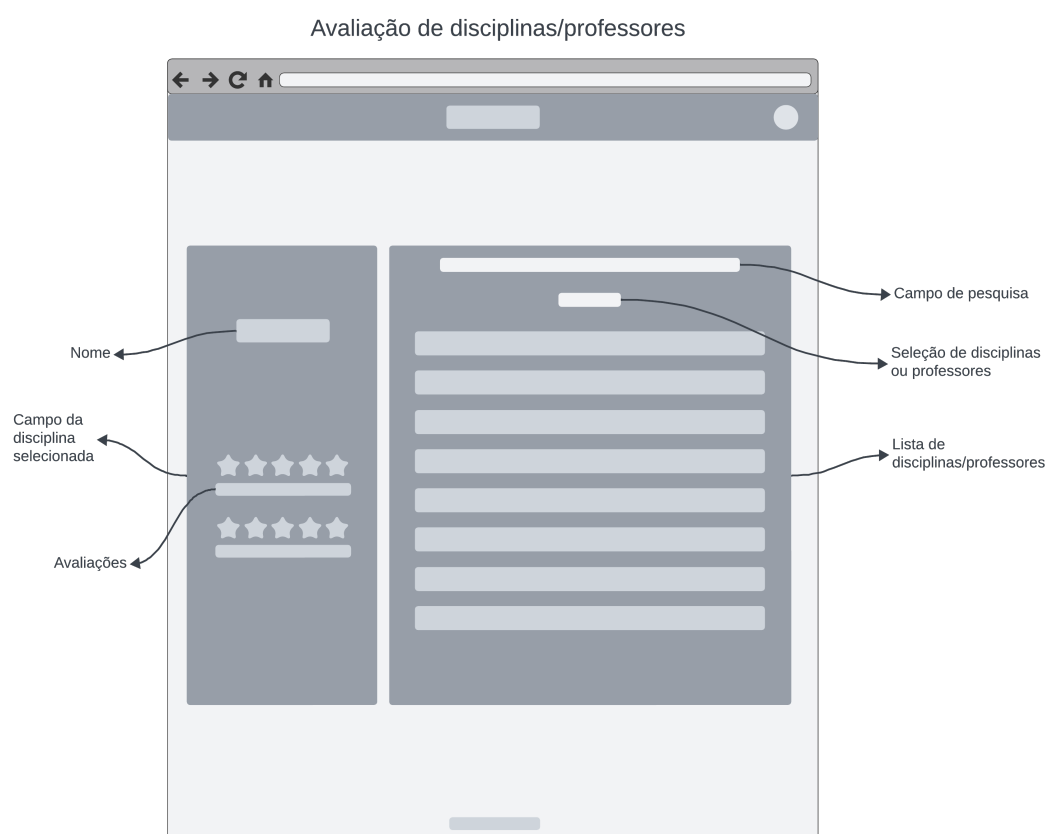


Figura 8.4: Wireframe da página de avaliação de disciplinas e professores

9

Dados do sistema

Para satisfazer os requisitos e os casos de uso, foi necessário modelar os dados disponíveis ao sistema e ao algoritmo.

9.1

Diagrama de entidades e relacionamentos

O objeto central dos dados do sistema é a *disciplina*. Uma disciplina tem alguns atributos, como código, nome, ementa e quantidade de créditos. Ela está contida em um *currículo*, é oferecida por um *departamento*, pode possuir como pré-requisito várias disciplinas, e pode possuir algumas *turmas*.

Cada turma pode possuir horários *horários*, pode conter várias *vagas* disponíveis a um destino, e é lecionada por um *professor*.

Cada professor tem como atributo um nome, e pode possuir diversas *avaliações* feitas por um *usuário*.

O usuário cursa um *currículo*, pode ter cursado múltiplas *disciplinas*, pode avaliar muitas *disciplinas* e *professores*, e pode possuir *grades* montadas no sistema.

Portanto, foram identificados nove entidades: **Disciplina**, **Departamento**, **Turma**, **Horário**, **Vagas**, **Professor**, **Usuário**, **Currículo** e **Grade**.

A figura 9.1 exibe o diagrama de entidades e relacionamentos (ER), que descreve os dados no modelos de entidades e seus relacionamentos com outras entidades.

Neste diagrama, um retângulo representa um entidade, e um retângulo com borda dupla representa uma entidade fraca, que é uma entidade que depende de outra para existir. Um losângulo representa um relacionamento entre duas entidades, e um losângulo com borda dupla representa um relacionamento de dependência de uma entidade fraca. Os atributos de um relacionamento são representados pelos círculos ovais. Um atributo que identifica uma entidade possui seu texto sublinhado. Um atributo que identifica uma entidade fraca (em conjunto com os atributos da entidade na qual depende para existir) tem seu texto com um sublinhado tracejado.

Este diagrama segue a notação de Peter Chen [8].

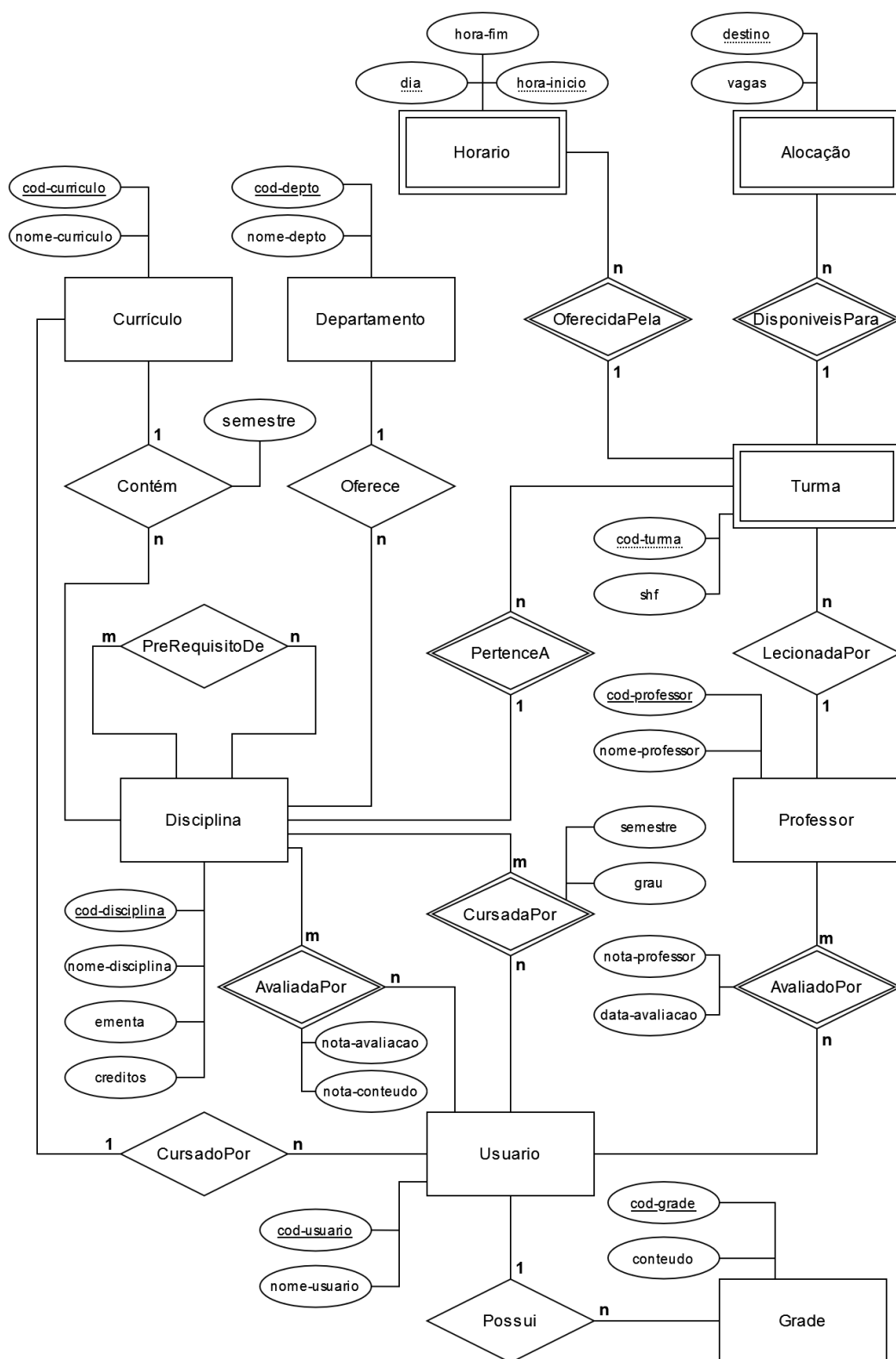


Figura 9.1: Diagrama de entidades e relacionamentos

9.2

Modelo lógico

A imagem 9.2 exibe o diagrama lógico dos dados com base no modelo de entidades e relacionamentos. Ele representa a estrutura implementada no

banco de dados, com suas tabelas e chaves primárias (PK) e chaves estrangeiras (FK).

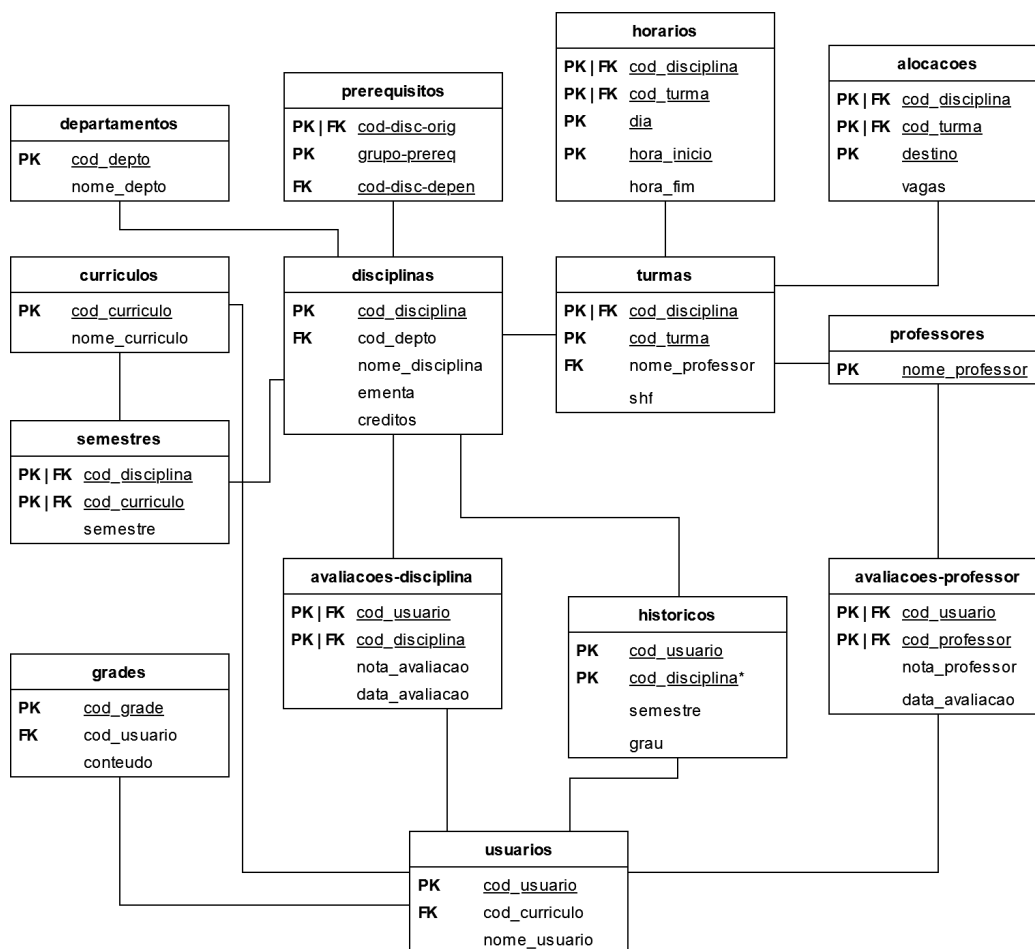


Figura 9.2: Diagrama do modelo lógico dos dados

9.3

Dicionário de dados

O dicionário de dados possui a função de categorizar os dados. Ele é uma coleção de metadados dos modelos conceitual e lógico. O dicionário de dados do sistema está disponível na tabela 9.1. Nele estão representadas todas as colunas observadas no diagrama lógico da figura 9.2. O tipo do dado é único para o nome da coluna, o que significa que o mesmo nome da coluna em duas tabelas diferentes representa o mesmo tipo de dado. Por exemplo, o tipo de dado na coluna *cod-usuario* é o mesmo na tabela *grades*, *usuarios*, *avaliacoes-disciplina* e *avaliacoes-professor*.

| Coluna | Descrição | Tipo | Domínio |
|----------------|--|------------------|---|
| cod_curriculo | Identificador do currículo. Exemplo: "CEGBCO20180" | string | Sem restrição. |
| cod_depto | Abreviação de três letras do departamento, conforme disponibilizado no microhorario. Exemplo: "ENG". | string | Três letras maiúsculas. |
| cod_disciplina | Código da disciplina. Exemplo: "INF1011". | string | Três letras maiúsculas (não necessariamente um departamento) seguidas de 4 dígitos. |
| cod_disc_depen | Códigos da disciplinas que fazem parte de um grupo de pré-requisitos de outra disciplina. | Array de strings | Mesmo do cod-disciplina. |
| cod_disc_orig | Código da disciplina que possui algum pré-requisito. | string | Mesmo do cod-disciplina. |
| cod_grade | Código da grade, gerado ao salvar uma nova grade. | string | 8 caracteres em formato base64. |
| cod_turma | Código da turma. | string | 3 caracteres. |
| cod_usuario | Código do usuário, representado por sua matrícula | string | 7 dígitos. |
| conteudo | Conteúdo codificado no formato JSON da grade horária, que é decodificado pela interface ao exibir. | string | Sem restrição. |

| Coluna | Descrição | Tipo | Domínio |
|----------------|---|--------|---|
| creditos | Quantidade de créditos da disciplina. Caso seja -1, então a disciplina não disponibilizou a quantidade de créditos. | int | Valor entre 0 a 16, ou -1. |
| data_avaliacao | Data em que a avaliação foi efetuada pelo aluno. | date | Uma data anterior à data atual. Pode ser nulo. |
| dia | Dia da semana que a turma é oferecida. | string | "SEG", "TER", "QUA", "QUI", "SEX", "SAB" ou "xxx" (sem dia definido). |
| destino | Destino das vagas disponíveis. Exemplo: "QQC" (Qualquer curso), "BCO" (Bacharelado em Engenharia de Computação), entre outros. | string | Três letras maiúsculas. Pode ser nulo. |
| ementa | Ementa da disciplina. | string | Sem restrição. Pode ser nulo. |
| grau | Grau do aluno na disciplina. | int | Entre 0 a 100. Pode ser nulo. |
| grupo_prereq | Um identificador do grupo do pré-requisito. Um aluno só pode cursar a disciplina se todas as disciplinas do mesmo grupo foram cursadas. | int | Um inteiro maior que zero. |

| Coluna | Descrição | Tipo | Domínio |
|-----------------|---|--------|--|
| hora_fim | Hora do fim da aula da turma. Se for nulo, então a turma não possui hora de fim. | int | Valor entre 9 e 23, deve ser maior que hora-inicio. Pode ser nulo. |
| hora_inicio | Hora do inicio da aula da turma. Se for zero (não pode ser nulo pois faz parte de uma chave primária), então a turma não possui hora de início. | int | Valor entre 7 e 21, ou 0. |
| nome_curriculo | Nome curto do currículo. Exemplo: "Engenharia 19.0". | string | Sem restrição. |
| nome_depto | Nome do departamento, conforme disponibilizado no microhorario. Exemplo: "Engenharia". | string | Sem restrição. |
| nome_disciplina | Nome da disciplina. Exemplo: "Semântica de Linguagens". | string | Sem restrição. |
| nome_professor | Nome do professor, conforme disponibilizado no microhorário | string | Sem restrição. |
| nome_usuario | Nome do usuário | string | Sem restrição. |
| nota_avaliacao | Nota da disciplina ou professor avaliado. | int | Valor entre 1 e 5. |
| semestre | Semestre recomendado para cursar a disciplina de acordo com o currículo. | int | Valor entre 1 e 10. Pode ser nulo. |

| Coluna | Descrição | Tipo | Domínio |
|--------|--|------|---|
| shf | Quantidade de horas "Sem Horário Fixo" de uma disciplina. | int | Valor maior ou igual a zero. Pode ser nulo. |
| vagas | Quantidade de vagas disponíveis. Se for -1, então não foi disponibilizada a quantidade de vagas para a turma | int | Inteiro maior ou igual a -1. |

Tabela 9.1: Dicionário de dados

10

Construção

Com a modelagem da interface e dos dados, foi possível desenvolver o algoritmo e o sistema. O desenvolvimento foi dividido em partes: definição das tecnologias utilizadas, método de coleta dos dados, definição e implementação do algoritmo, e implementação do sistema (interface, API e microserviços). Cada uma das partes está explicada e descrita nas seções a seguir.

10.1

Tecnologias utilizadas

Para o armazenamento e consulta dos dados, foi utilizado o banco de dados relacional PostgreSQL [9]. O PostgreSQL foi escolhido por ser robusto e eficiente, mas fácil e rápido de configurar.

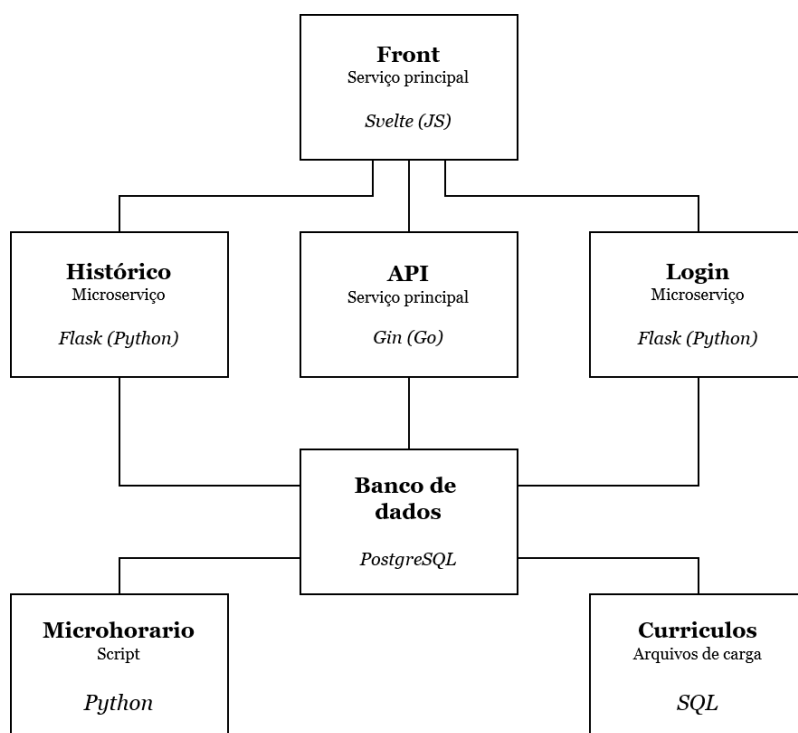


Figura 10.1: Diagrama da arquitetura do sistema

A figura 10.1 exibe um diagrama da arquitetura do sistema. O sistema foi separado em interface e API (*Application programming interface*, ou interface de programação da aplicação). A coleta dos dados provenientes do Microhorario, a autenticação com o Sistema Acadêmico Universitário e carga/processamento dos históricos foi desenvolvida em serviços menores uma única funcionalidade respectivamente (microserviços).

As tecnologias utilizadas na API, nos microserviços e na interface estão descritas respectivamente nas seções 10.4, 10.5 e 10.6.

10.2

Coleta de dados

O modelo lógico representado na figura 9.2 define a construção de catorze tabelas no banco de dados. O trecho de código 1 contém o modelo físico do banco de dados, ou seja o código SQL que define a criação das tabelas.

Código 1: Modelo físico

```

1 CREATE TABLE IF NOT EXISTS departamentos (
2     cod_depto CHAR(3) PRIMARY KEY,
3     nome_depto TEXT NOT NULL
4 );
5
6 CREATE TABLE IF NOT EXISTS disciplinas (
7     cod_disciplina CHAR(7) PRIMARY KEY,
8     cod_depto CHAR(3) NOT NULL,
9     nome_disciplina TEXT NOT NULL,
10    ementa TEXT,
11    credits SMALLINT NOT NULL CHECK (credits >= -1 AND credits <=
12        16),
13    FOREIGN KEY (cod_depto) REFERENCES departamentos(cod_depto)
14 );
15
16 CREATE TABLE IF NOT EXISTS prerequisitos (
17     cod_disc_orig CHAR(7) NOT NULL,
18     grupo_prereq SMALLINT NOT NULL,
19     cod_disc_depen CHAR(7) NOT NULL,
20     PRIMARY KEY (cod_disc_orig, grupo_prereq, cod_disc_depen),
21     FOREIGN KEY (cod_disc_orig) REFERENCES disciplinas(
22         cod_disciplina)
23 );
24
25 CREATE TABLE IF NOT EXISTS professores (
26     nome_professor TEXT PRIMARY KEY
27 );
28
29 CREATE TABLE IF NOT EXISTS turmas (
30     cod_turma CHAR(3) NOT NULL,
31     cod_disciplina CHAR(7) NOT NULL,
32     nome_professor TEXT NOT NULL,
33     shf SMALLINT NOT NULL,
34     PRIMARY KEY (cod_turma, cod_disciplina),
35     FOREIGN KEY (cod_disciplina) REFERENCES disciplinas(
36         cod_disciplina),
37     FOREIGN KEY (nome_professor) REFERENCES professores(
38         nome_professor)
39 );

```

```
36
37 CREATE TABLE IF NOT EXISTS horarios (
38     cod_disciplina CHAR(7) NOT NULL,
39     cod_turma CHAR(3) NOT NULL,
40     dia CHAR(3) NOT NULL,
41     hora_inicio SMALLINT NOT NULL CHECK (hora_inicio = 0 OR (
42         hora_inicio >= 7 AND hora_inicio <= 21)),
43     hora_fim SMALLINT,
44     PRIMARY KEY (cod_disciplina, cod_turma, dia, hora_inicio),
45     FOREIGN KEY (cod_disciplina, cod_turma) REFERENCES turmas(
46         cod_disciplina, cod_turma)
47 );
48
49 CREATE TABLE IF NOT EXISTS alocacoes (
50     cod_disciplina CHAR(7) NOT NULL,
51     cod_turma CHAR(3) NOT NULL,
52     destino VARCHAR(80) NOT NULL,
53     vagas SMALLINT NOT NULL CHECK (vagas >= -1),
54     PRIMARY KEY (cod_disciplina, cod_turma, destino),
55     FOREIGN KEY (cod_disciplina, cod_turma) REFERENCES turmas(
56         cod_disciplina, cod_turma)
57 );
58
59 CREATE TABLE IF NOT EXISTS curriculos (
60     cod_curriculo VARCHAR(11) PRIMARY KEY,
61     nome_curriculo TEXT NOT NULL
62 );
63
64 CREATE TABLE IF NOT EXISTS semestres (
65     cod_disciplina CHAR(7) NOT NULL,
66     cod_curriculo VARCHAR(11) NOT NULL,
67     semestre SMALLINT CHECK (semestre >= 0 AND semestre <= 10),
68     PRIMARY KEY (cod_disciplina, cod_curriculo),
69     FOREIGN KEY (cod_disciplina) REFERENCES disciplinas(
70         cod_disciplina),
71     FOREIGN KEY (cod_curriculo) REFERENCES curriculos(cod_curriculo)
72 );
73
74 CREATE TABLE IF NOT EXISTS usuarios (
75     cod_usuario CHAR(7) PRIMARY KEY,
76     nome_usuario TEXT NOT NULL,
77     cod_curriculo VARCHAR(11), -- inicialmente o usuario nao
78         cadastrou o curriculo
79     FOREIGN KEY (cod_curriculo) REFERENCES curriculos(cod_curriculo)
80 );
81
82 CREATE TABLE IF NOT EXISTS grades (
83     cod_grade CHAR(8) PRIMARY KEY,
84     cod_usuario CHAR(7) NOT NULL,
85     conteudo TEXT NOT NULL,
86     FOREIGN KEY (cod_usuario) REFERENCES usuarios(cod_usuario)
```



```

82 );
83
84 CREATE TABLE IF NOT EXISTS historicos (
85     cod_usuario CHAR(7) NOT NULL,
86     cod_disciplina CHAR(7) NOT NULL,
87     semestre SMALLINT NOT NULL,
88     grau SMALLINT, -- a disciplina pode nao ter nota
89     PRIMARY KEY (cod_usuario, cod_disciplina, semestre),
90     FOREIGN KEY (cod_usuario) REFERENCES usuarios(cod_usuario),
91     FOREIGN KEY (cod_disciplina) REFERENCES disciplinas(cod_usuario)
92 );
93
94 CREATE TABLE IF NOT EXISTS avaliacoes_disciplinas (
95     cod_usuario CHAR(7) NOT NULL,
96     cod_disciplina CHAR(7) NOT NULL,
97     nota_avaliacao SMALLINT NOT NULL CHECK (nota_avaliacao >= 0 AND
98         nota_avaliacao <= 5),
99     data_avaliacao DATE NOT NULL,
100    PRIMARY KEY (cod_usuario, cod_disciplina),
101    FOREIGN KEY (cod_usuario) REFERENCES usuarios(cod_usuario)
102    FOREIGN KEY (cod_disciplina) REFERENCES disciplinas(
103        cod_disciplina)
104 );
105
106 CREATE TABLE IF NOT EXISTS avaliacoes_professores (
107     cod_usuario CHAR(7) NOT NULL,
108     nome_professor TEXT NOT NULL,
109     nota_avaliacao SMALLINT NOT NULL CHECK (nota_avaliacao >= 0 AND
110         nota_avaliacao <= 5),
111     data_avaliacao DATE NOT NULL,
112     PRIMARY KEY (cod_usuario, nome_professor),
113     FOREIGN KEY (cod_usuario) REFERENCES usuarios(cod_usuario)
114     FOREIGN KEY (nome_professor) REFERENCES professores(
115         nome_professor)
116 );
117
118 CREATE TABLE IF NOT EXISTS modificacao (
119     data_ementa timestamptz NOT NULL,
120     data_geral timestamptz NOT NULL,
121     modo_fallback boolean NOT NULL,
122     PRIMARY KEY (data_ementa, data_geral, modo_fallback)
123 )

```

No modelo físico, é possível observar que as avaliações, históricos e currículos devem ser sempre associados às disciplinas, através das chaves estrangeiras. Portanto, as disciplinas não podem ser excluídas, mesmo que elas possam não estar sendo oferecidas no semestre atual. Para diferenciar uma disciplina que está sendo oferecida de uma disciplina que não está disponível,

basta verificar se ela existe na tabela `turmas`.

No modelo físico, existe a criação de uma tabela `modificacao`, que não estava prevista no modelo lógico. Essa tabela não contém dados relacionais, mas apenas três valores armazenados em uma só linha da tabela. Esses valores representam as datas de coleta das informações das disciplinas e turmas, e se os dados estão em modo de `fallback`, que será explicado na seção 10.2.1.

Além dos dados gerados pela interação do usuário, o sistema precisa de dados provenientes da faculdade, como as disciplinas, turmas, professores e currículos. Esses dados são coletados ocasionalmente, exceto os currículos, que foram manualmente construídos a partir de informações disponíveis nas plataformas da universidade. Como o algoritmo e o sistema foram planejado para alunos do Departamento de Informática, a quantidade de currículos a ser inserida é pequena. No total, foram inseridos seis currículos: três de engenharia de computação (Currículos 2023, 2018.0 e 2018.1) e três de ciência de computação, disponíveis nas respectivas páginas ^{1 2} dos cursos.

Os dados de cada currículo foram transformados manualmente em um arquivo no formato SQL que insere o código do currículo e constrói uma `PROCEDURE`, um procedimento responsável por inserir as disciplinas referentes ao currículo. As disciplinas estão dentro de uma `procedure` pois esta deve ser executada toda vez que uma disciplina nova possa ter sido inserida no banco. Isso é necessário pois a `procedure` armazena somente o código de disciplinas que existem no banco, devido a restrição de chave estrangeira da tabela. Portanto, cada vez que uma disciplina nova é inserida, essa operação precisa ser refeita. O trecho de código 2 contém um exemplo de um dos seis arquivos criados, um para cada currículo.

Código 2: Exemplo do arquivo de carga de currículos

```

1 INSERT INTO curriculos (cod_curriculo, nome_curriculo)
2 SELECT * FROM (
3     VALUES ('CCPBCC20181', 'Ciência da Computação (Bacharelado) -
4             Currículo 18.1')
5 ) AS i (cod_curriculo, nome_curriculo)
6 WHERE NOT EXISTS (
7     SELECT cod_curriculo
8     FROM curriculos
9     WHERE cod_curriculo = i.cod_curriculo
10 );
11 CREATE OR REPLACE PROCEDURE inserir_curriculo_cie_18_1()
12 LANGUAGE SQL
13 BEGIN ATOMIC

```

¹https://www.puc-rio.br/ensinopesq/ccg/eng_computacao.html

²https://www.puc-rio.br/ensinopesq/ccg/ciencia_computacao.html

```

14      DELETE FROM semestres WHERE cod_curriculo = 'CCPBCC20181';
15
16      INSERT INTO semestres (cod_curriculo, cod_disciplina, semestre)
17      SELECT * FROM (VALUES
18          ('CCPBCC20181', 'INF1025', 1),
19          ('CCPBCC20181', 'INF1012', 1),
20          ('CCPBCC20181', 'INF1009', 1),
21          ('CCPBCC20181', 'INF1031', 1),
22          -- ....
23          -- ....
24          ('CCPBCC20181', 'INF0381', 8),
25          ('CCPBCC20181', 'INF1951', 8)
26      ) AS i (cod_curriculo, cod_disciplina, semestre)
27      WHERE EXISTS (
28          SELECT cod_disciplina
29          FROM disciplinas
30          WHERE cod_disciplina = i.cod_disciplina
31      );
32 END

```

Os dados provenientes das disciplinas e turmas são coletados diretamente através do microhorário. Foi criada uma biblioteca Python que permite baixar todos as informações disponíveis no microhorário, e opcionalmente coletar as ementas e pré-requisitos que estão disponíveis em página. A biblioteca baixa os dados das disciplinas e turmas em menos de três segundos, mas pode executar por até 30 minutos para baixar as ementas e pré-requisitos que estão em outro serviço da universidade. Portanto, foi criado um módulo, chamado `microhorario`, que permite atualizar o banco com as informações básicas das disciplinas e turmas, mas que pode ser alterado para atualizar também as ementas e pré-requisitos. Esse módulo foi configurado de forma a ser executado na sua forma mais simples (somente atualiza os dados referentes às turmas) uma vez a cada hora, e executado na forma completa (coleta inclusive as ementas e pré-requisitos) uma vez por dia.

10.2.1

Microhorário no fim do período

No final de cada período, devido a preparação das disciplinas e turmas para o próximo semestre, o microhorário não disponibiliza dados como quantidade de vagas e quantidade de créditos. Por isso, foi necessário que o sistema e o algoritmo acomodasse essas necessidades temporárias. Nesse caso, a variável de `fallback`, presente na tabela de `modificacao` estará no seu valor verdadeiro. Se o microhorário estiver disponível, essa variável está com seu valor falso.

10.2.2

Disciplinas optativas

Os currículos contém grupos de disciplinas optativas. Cada grupo de disciplinas optativas possuem várias disciplinas, e o aluno deve cursar pelo menos uma delas para que seja considerado o grupo como cursado. Por exemplo, o currículo de 19.0 de Engenharia de Computação determina o grupo de optativa INF0303, com as disciplinas ENG1456 - INTELIGENCIA COMPUT APLICADA e INF1771 - INTELIGENCIA ARTIFICIAL, e o aluno deve necessariamente cursar uma das duas. Essa categoria de disciplinas não foi percebida durante todo o processo de entrevistas, levantamento de requisitos e modelagem do banco. Para implementar essa categoria de forma eficaz, seria necessária uma reestruturação das tabelas do banco de dados, e diversas alterações complexas nos módulos de coleta de dados do microhorário e dos outros serviços, inclusive alterações na biblioteca de consulta ao microhorário citada acima. Portanto, optou-se por não implementar as disciplinas optativas. Em vez disso, as disciplinas relacionadas ao grupo de optativas não são atreladas ao currículo, sendo consideradas eletivas.

10.2.3

Co-requisitos

Durante o desenvolvimento de coleta de dados da ementa, foi observado que algumas disciplinas possuem co-requisitos. Co-requisitos são disciplinas que precisam ser cursadas no mesmo período, conforme a imagem 10.2. Como esse atributo foi descoberto no momento que o desenvolvimento já estava em andamento, optou-se por não implementar essa funcionalidade. Em vez disso, nenhuma disciplina possui co-requisitos.

10.2.4

Quantidade de créditos

Durante a fase de entrevistas, foi solicitado a média de créditos das grades horários dos entrevistados. O objetivo era possivelmente alimentar o algoritmo de recomendação com essa média para, por exemplo, recomendar disciplinas com menos créditos caso o aluno estivesse chegando nessa média. Porém, não foi possível criar um bom modelo matemático utilizando essa informação, por isso ela foi descartada na implementação atual.



Figura 10.2: Seleção de uma disciplina com co-requisitos durante o processo de matrícula

10.3

Definição e implementação do algoritmo

O algoritmo busca satisfazer as necessidades apresentadas na tabela 5.3. As entrevistas indicaram a importância de cinco categorias (Conteúdo, Professor, Avaliação, Horário e Opinião). Por isso, o algoritmo foi dividido nessas cinco categorias.

O algoritmo se baseia num valor $V[d, u]$ atribuído a cada disciplina, que indica a relevância da disciplina d para o usuário u . O valor $V[d, u] \in [0, 1]$, sendo 1 o maior grau de relevância da disciplina, e 0 o menor. O cálculo de V é uma média ponderada, conforme a equação 10-1.

$$V[d, u] = P_c V_c[d, u] + P_p V_p[d] + P_o V_o[d] + P_h V_h[d, u] + P_a V_a[d] \quad (10-1)$$

Em que d é uma disciplina, u é um usuário, V_x é o valor de uma das cinco categorias para a disciplina d e o usuário u , e P_x é o peso de uma das cinco categorias. Por exemplo, $P_c V_c[d, u]$ se refere ao cálculo do conteúdo, $P_p V_p[d]$ se refere ao cálculo do professor, e assim por diante.

O valor $V_c[d, u]$ indica o quão relevante é a disciplina para o usuário de acordo com o seu conteúdo. Para isso, o valor é calculado de acordo com o

histórico de outros alunos e com o currículo do aluno, conforme a equação 10-2.

$$V_c[d, u] = \begin{cases} 1.0 & \text{caso } d \in \text{Historico}[u] \\ \frac{|A_{cursou}[d, u]|}{|A_{curriculo}[u]|} & \text{caso } A_{curriculo}[u] > 0 \\ 0 & \text{caso contrário} \end{cases} \quad (10-2)$$

Em que:

$$\begin{aligned} A_{curriculo}[u] &= \forall a \in \text{Alunos} \mid \text{Curriculo}[a] = \text{Curriculo}[u] \\ A_{cursou}[d, u] &= \forall a \in A_{curriculo}[u] \mid d \in \text{Cursadas}[a] \end{aligned}$$

Em que d é uma disciplina, u é um usuário, Alunos é o conjunto de todos os usuários cadastrados no sistema, $\text{Curriculo}[u]$ é o currículo do usuário u , e $\text{Cursadas}[u]$ é o conjunto de disciplinas que o aluno u cursou. Em resumo, o valor $V_c[d, u]$ é o valor máximo caso a disciplina deve ser cursada pelo usuário, ou sendo uma eletiva é a proporção de alunos do mesmo currículo que fizeram esta disciplina. Essa proporção indica se o conteúdo é relevante para alunos semelhantes ao usuário.

O valor $V_p[d]$ indica o quão relevante é a disciplina para o usuário de acordo com o professor. Para isso, o valor é calculado de acordo com as avaliações dos professores das turmas das disciplinas, conforme a equação 10-3.

$$V_p[d] = \frac{\sum_{p \in P[d]} \frac{\sum_{a \in \text{Avs}[p]} a}{|\text{Avs}[p]|}}{|P[d]|} / 5 \quad (10-3)$$

Em que d é uma disciplina, $P[d]$ é o conjunto dos professores que lecionam a disciplina d , e $\text{Avs}[p]$ representa o conjunto das avaliações dos usuários do professor p . Em resumo, o valor $V_p[d]$ é a média das avaliações de todos os professores que estão lecionando a disciplina, com o valor entre 0 e 1.

O valor $V_o[d]$ indica o quão relevante é a disciplina para o usuário de acordo com a opinião dos alunos. A equação é semelhante ao cálculo apresentado na equação 10-3. Porém, nesse caso, é levado em conta a média das avaliações das próprias disciplinas, conforme a equação 10-4.

$$V_o[d] = \frac{\sum_{a \in \text{Op}[d]} a}{|\text{Op}[d]|} / 5 \quad (10-4)$$

Em que d é uma disciplina, e $\text{Op}[d]$ é o conjunto das avaliações dos usuários da disciplina d .

O valor $V_h[d, u]$ indica o quão relevante é a disciplina para o aluno de acordo com o horário. Para isso, o valor é calculado de acordo com os horários ocupados na grade do usuário e os horários das disciplinas, conforme a equação 10-5

$$V_h[d, u] = \begin{cases} \frac{|T_{possiveis}[d, u]|}{|T[d]|} & \text{caso } T[d] > 0 \\ 0 & \text{caso contrário} \end{cases} \quad (10-5)$$

Em que d é uma disciplina, u é um usuário, $T_{possiveis}[d, u]$ é o conjunto de turmas da disciplina d que se encaixam na grade do usuário u , e $T[d]$ é a quantidade de turmas da disciplina d . Em resumo, $V_h[d, u]$ é a porcentagem de turmas da disciplina que se encaixam na grade do usuário.

Por último, o valor $V_a[d]$ indica o quão relevante é a disciplina para o aluno de acordo com os graus (a nota final, considerando as provas do aluno durante o semestre) da disciplina. O cálculo é semelhante às equações 10-3 e 10-4, conforme a equação 10-6.

$$V_p[d] = \frac{\sum_{g \in Graus[d]} g}{|Graus[d]|} / 100 \quad (10-6)$$

Em que d é uma disciplina e $Graus[d]$ é o conjunto das notas da disciplina d .

Para obter todos os valores dos conjuntos citados nas equações anteriores, foi criada uma consulta SQL que obtém todos os valores de uma só vez.

O trecho de código 3 contém a consulta por completo. Antes de executar, são substituídos os valores `@usuario` pelo código do usuário, e o valor `@escolhas` por um trecho de código SQL referente às escolhas de turmas de disciplinas na grade atual do aluno.

Código 3: Consulta dos dados para o algoritmo

```

1 WITH rec_c1 AS (
2     SELECT cod_disciplina
3     FROM semestres
4     WHERE cod_curriculo =
5         (SELECT cod_curriculo FROM usuarios where cod_usuario =
6           @usuario)
7 ), rec_c2_1 AS (
8     SELECT s.cod_disciplina, count(u.cod_usuario) as qtd
9     FROM semestres s, usuarios u
10    WHERE s.cod_curriculo = u.cod_curriculo
11    AND u.cod_curriculo =
12        (SELECT cod_curriculo FROM usuarios where cod_usuario =
13          @usuario)
14    GROUP BY s.cod_disciplina
15 ), rec_c2_2 AS (

```

```

14     SELECT count(cod_usuario)
15     FROM usuarios
16     WHERE cod_curriculo =
17           (SELECT cod_curriculo FROM usuarios where cod_usuario =
18             @usuario)
19 ), rec_h AS (
20     SELECT h1.cod_disciplina, count(DISTINCT h1.cod_turma) as
21       possiveis, todas
22     FROM horarios h1
23     LEFT JOIN (
24       SELECT cod_disciplina, count(cod_turma) as todas
25       FROM turmas
26       GROUP BY cod_disciplina
27     ) AS h2 ON h1.cod_disciplina = h2.cod_disciplina
28     WHERE (dia, hora_inicio, hora_fim) NOT IN (
29       SELECT dia, hora_inicio, hora_fim
30       FROM horarios
31       WHERE @escolhas
32     )
33     GROUP BY h1.cod_disciplina, todas
34 ), rec_o AS (
35     SELECT cod_disciplina, avg(nota_avaliacao) as media
36     FROM avaliacoes_disciplinas
37     GROUP BY cod_disciplina
38 ), rec_p AS (
39     SELECT d.cod_disciplina, avg(a.nota_avaliacao) as media
40     FROM disciplinas d, avaliacoes_professores a, turmas t
41     WHERE d.cod_disciplina = t.cod_disciplina
42     AND t.nome_professor = a.nome_professor
43     GROUP BY d.cod_disciplina, d.nome_disciplina
44 ), rec_a AS (
45     SELECT cod_disciplina, avg(grau) as media
46     FROM historicos
47     GROUP BY cod_disciplina
48 ), filtro AS (
49     SELECT d.cod_disciplina
50     FROM disciplinas d
51     WHERE cod_disciplina NOT IN (
52       SELECT h.cod_disciplina
53       FROM historicos h
54       WHERE cod_usuario = @usuario
55     )
56 )
57 SELECT f.cod_disciplina,
58       c1.cod_disciplina IS NOT NULL conteudo1,
59       c2_1.qtd conteudo21,
60       c2_2.count conteudo22,
61       h.cod_disciplina IS NOT NULL horario,
62       o.media opiniao,
63       p.media professor,

```



```

63      a.media avaliacao
64 FROM filtro f
65      LEFT JOIN rec_c1 c1 ON f.cod_disciplina = c1.cod_disciplina
66      LEFT JOIN rec_c2_1 c2_1 ON f.cod_disciplina = c2_1.
        cod_disciplina
67      LEFT JOIN rec_h h ON f.cod_disciplina = h.cod_disciplina
68      LEFT JOIN rec_o o ON f.cod_disciplina = o.cod_disciplina
69      LEFT JOIN rec_p p ON f.cod_disciplina = p.cod_disciplina
70      LEFT JOIN rec_a a ON f.cod_disciplina = a.cod_disciplina,
71      rec_c2_2 c2_2

```

A consulta possui oito consultas auxiliares, sendo sete delas usadas para calcular os valores V_x do algoritmo de recomendação. As explicações das consultas auxiliares estão na lista a seguir:

- **rec_c1** Seleciona as disciplinas oferecidas pelo currículo do usuário;
- **rec_c2_1** Seleciona a quantidade de alunos do mesmo currículo do usuário por disciplina do currículo do usuário;
- **rec_c2_2** Seleciona a quantidade de usuários do mesmo currículo do usuário;
- **rec_h** Seleciona, por disciplina, a quantidade de turmas não conflitantes com a grade do usuário e a quantidade total de turmas;
- **rec_o** Seleciona as média das avaliações das disciplinas;
- **rec_p** Seleciona as médias das avaliações dos professores;
- **rec_a** Seleciona as médias dos graus dos alunos por disciplina;
- **filtro** Seleciona as disciplinas que o usuário não cursou.

Por fim, os pesos P_x da média ponderada foram baseados na tabela 5.3. Cada peso é a soma dos pesos dos critérios, de acordo com as equações a seguir.

$$P_c = \frac{41}{41 + 19 + 8 + 28 + 21} \quad (10-7)$$

$$P_p = \frac{28}{41 + 19 + 8 + 28 + 21} \quad (10-8)$$

$$P_o = \frac{8}{41 + 19 + 8 + 28 + 21} \quad (10-9)$$

$$P_h = \frac{19}{41 + 19 + 8 + 28 + 21} \quad (10-10)$$

$$P_a = \frac{21}{41 + 19 + 8 + 28 + 21} \quad (10-11)$$

10.4

Implementação da API

A API é responsável por servir disponibilizar à interface os dados de forma organizada e eficaz. Foi estudada a implementação da API em quatro possíveis frameworks em três diferentes linguagens: *Django* [10], *Flask* [11], *Gin* [12] e *Rocket* [13]. Cada um dos frameworks possui suas vantagens e desvantagens.

O framework *Django* é um *Web Framework* completo, que possui múltiplas funcionalidades pré-configuradas, possui uma interface de comunicação com banco de dados bem robusta. O framework *Flask* por sua vez não possui muitas funcionalidades embutidas, e depende da instalação de pacotes externos para ampliar suas funcionalidades. Ambos os frameworks são desenvolvidos na linguagem Python, o que torna o desenvolvimento mais fácil, mas reduz performance do funcionamento, por ser uma linguagem interpretada.

O framework *Rocket* é desenvolvido na linguagem Rust [14], conhecida por ser rápida e segura, por possuir uma abordagem de manipulação de memória diferente de outras linguagens. Porém, Rust possui uma alta curva de aprendizado, dificultando o desenvolvimento do código.

Por fim, o framework *Gin* é desenvolvido na linguagem Go [15], conhecida por bem eficiente, útil para o desenvolvimento de APIs pelo sua capacidade de multiprocessamento, e fácil de usar. Por isso, esse framework foi escolhido para a API do sistema.

A API disponibiliza a documentação completa de todas as suas rotas, com as respectivas entradas e saídas, conforme a figura 10.3.

10.5

Implementação dos microserviços

Os serviços de autenticação, e de carga e processamento dos históricos dos alunos são processos separados da API. Ambos foram desenvolvidos na linguagem Python, por sua facilidade de processamento dos dados, e por não haver a necessidade uma performance ótima. Esses utilizam o framework Flask para disponibilizar os serviços, por serem serviços bem simples.

O serviço de autenticação se comunica com o Sistema Acadêmico Universitário (SAU). Existe uma API para autenticação dos alunos da universidade, mas é restrita para os serviços interno da mesma. Por isso, o serviço de autenticação implementado simula um aluno autenticando-se no portal ³ do SAU, e verifica se a autenticação foi efetuada com sucesso.

³Disponível em: <<https://www.puc-rio.br/ensinopesq/academicas/>>, Sistemas Acadêmicos - SAU

O serviço de carga dos históricos precisa receber como entrada a página do histórico do usuário. É possível coletar o histórico de forma automática, acessando o portal do SAU também simulando um usuário. Porém, como essa interação com o SAU não seria transparente com o usuário, sendo executado de forma oculta, não foi considerada a melhor opção. Por isso, optou-se pelo usuário salvando uma cópia da página do seu histórico, e submetendo-o através da interface, que se comunica com o serviço de carga e processamento dos históricos.

10.6

Implementação da interface

A interface foi implementada utilizando o framework *Svelte*, que permite desenvolver páginas interativas baseadas em componentes. A programação é feita utilizando um misto de Javascript e HTML, que é compilada em pacotes Javascript pequenos que são interpretados pelo navegador de internet.

Como foi necessário desenvolver pelo menos três páginas diferentes, conforme o wireframe do capítulo 8, foi utilizado o framework *SvelteKit* [16]. Este permite disponibilizar páginas desenvolvidas em Svelte em diferentes rotas, além de outras possibilidades, como SSR (*Server-Side Rendering*, ou renderização no servidor), que permite reduzir o esforço do navegador do usuário ao construir inicialmente a página no servidor.

A seguir estão algumas imagens da interface implementada do sistema. A figura 10.4 mostra a tela inicial, com um menu de seleção para a página de grade e de avaliações. As figuras 10.7 e 10.8 exibem a interface de criação de grade horária. Na figura 10.7 é possível observar a mensagem de aviso, caso os dados completos do microhorário estejam indisponíveis.

As disciplinas recomendadas podem possuir informações extras. A figura 10.10 mostra uma legenda disponível no sistema explicando as informações exibidas. A figura explica que uma disciplina pode estar sendo recomendada devido a 5 categorias, que estão associadas aos pesos explicados na seção 10.3.

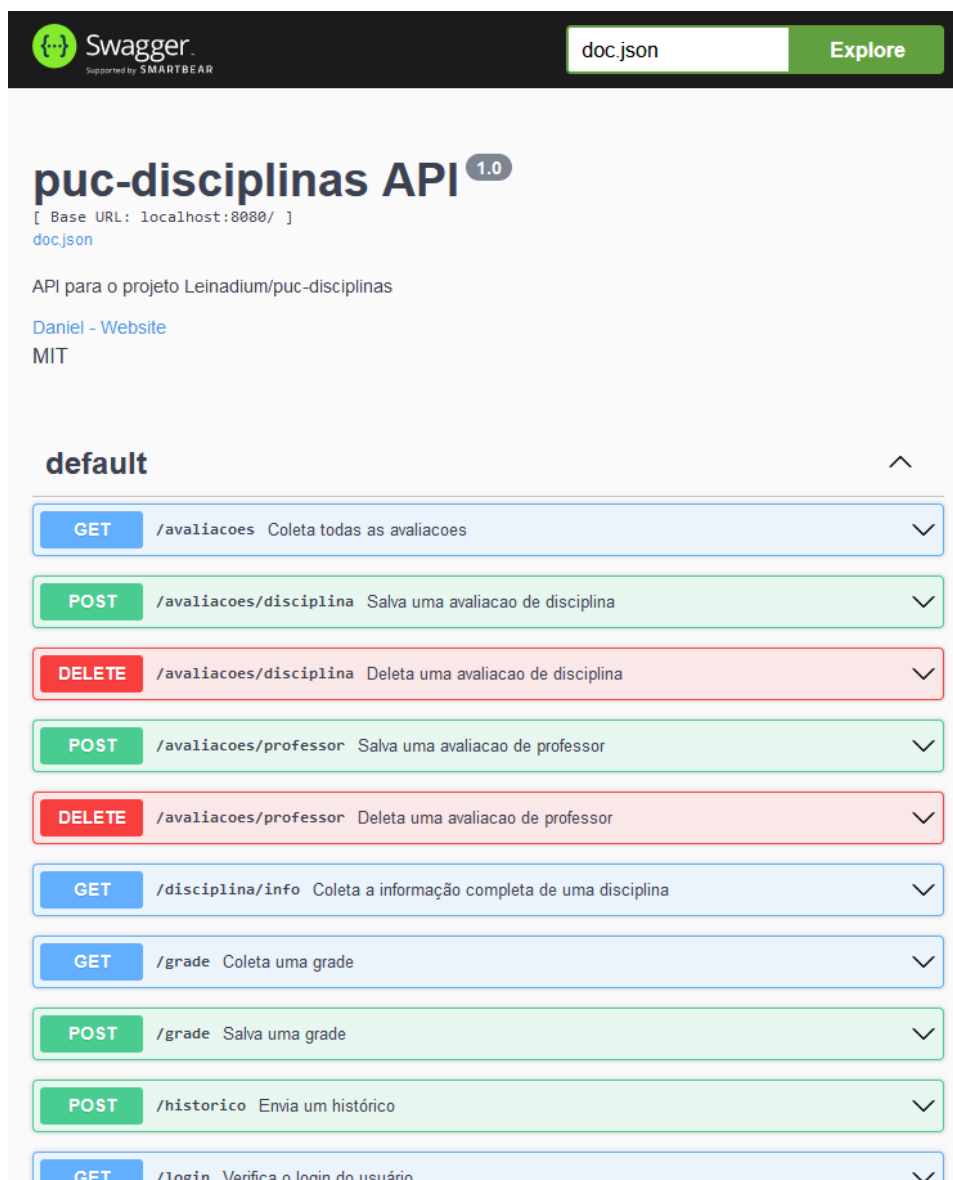


Figura 10.3: Documentação da API

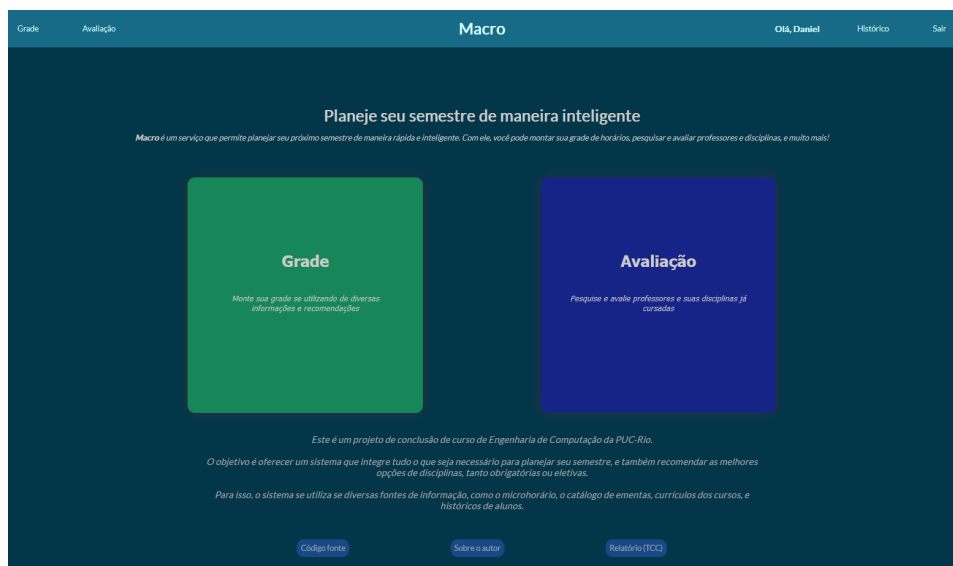


Figura 10.4: Implementação da tela inicial da interface

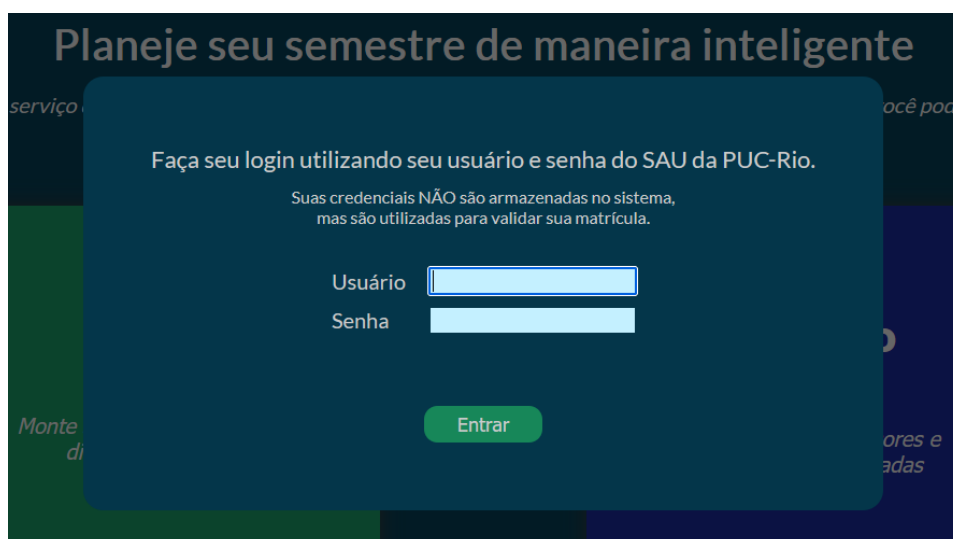


Figura 10.5: Implementação da tela de autenticação do usuário

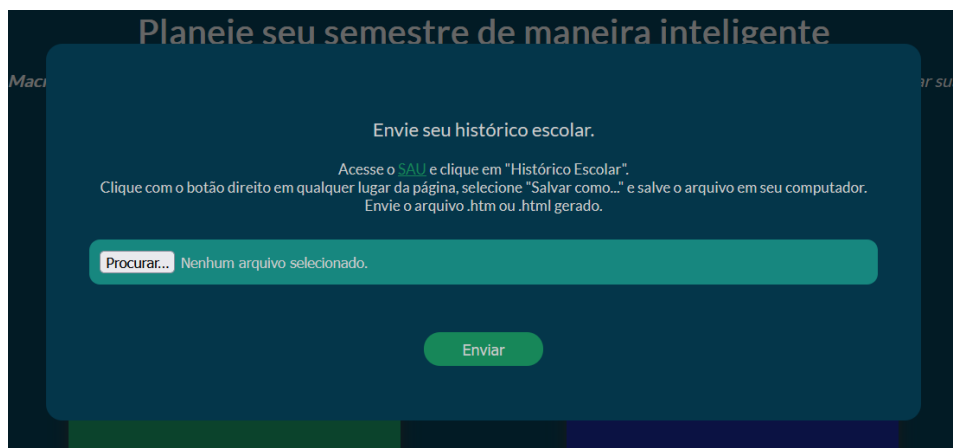


Figura 10.6: Implementação da tela de cadastro de histórico

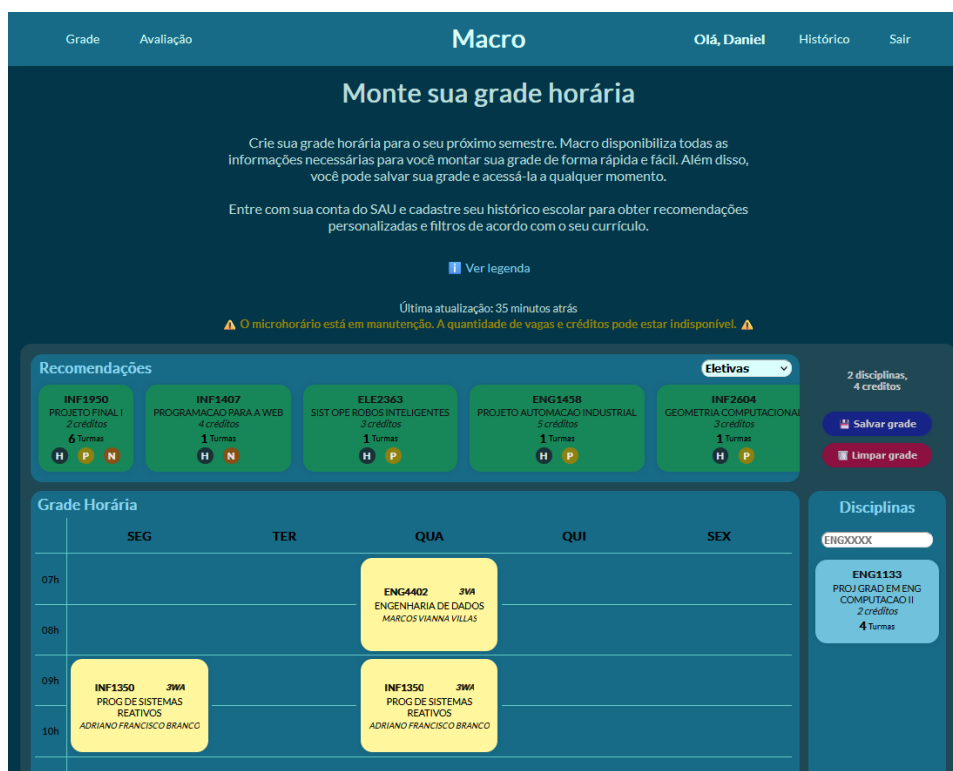


Figura 10.7: Implementação da tela de criação de grade da interface



Figura 10.8: Detalhes de uma disciplina na interface, exibido ao selecionar uma disciplina



Figura 10.9: Implementação da tela de avaliação da interface

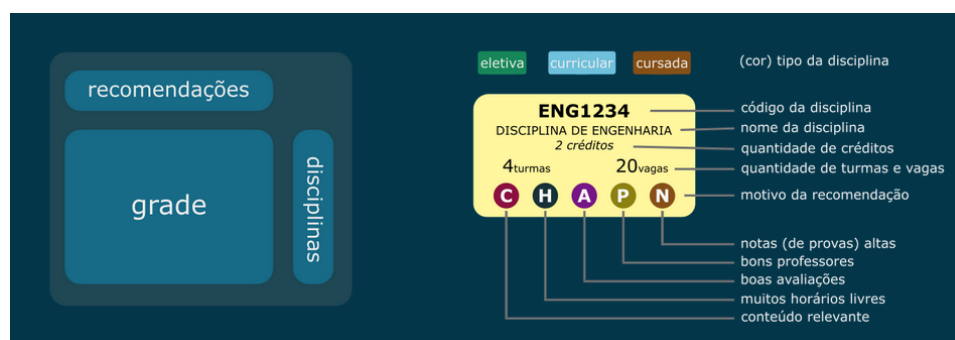


Figura 10.10: Legenda da interface de criação de grade horária

11

Testes

Afim de satisfazer todos os requisitos descritos no capítulo 6, foram efetuados testes do sistema e algoritmo durante toda a fase de desenvolvimento. Esses testes eram efetuados de maneira não estruturada, apenas para testar todos os possíveis casos de erro. Foi considerada a metodologia de *Desenvolvimento orientado a testes*, que planeja o desenvolvimento primeiro de testes unitários e de integração, para depois o desenvolvimento da aplicação em si. Porém, a implantação dessa metodologia costuma aumentar consideravelmente o tempo de desenvolvimento de aplicações, por isso ela não foi implementada.

Por não ter testes unitários ou de integração, o desenvolvimento precisou ser meticuloso. Por isso, optou-se por utilizar Typescript na interface, um superconjunto de Javascript, o que significa que suporta a sintaxe de Javascript e suas funcionalidades, mas introduz outras funcionalidades, que no caso são os tipos. Esses tipos permitem analisar o código em tempo de compilação, exibindo-os durante o processo de desenvolvimento, tornando o código mais resistente a falhas. Do mesmo modo, nos microserviços implementados em Python foram usados a sintaxe de tipos, pela mesma vantagem. Python é uma linguagem dinâmica em que os tipos não são necessários, mas foram utilizados para trazer mais segurança. Por último, a API foi implementada em Go, uma linguagem compilada *fortemente tipada*, o que significa que todas as variáveis possuem um tipo específico.

11.1

Cobertura dos requisitos

Antes de verificar o funcionamento do sistema e do algoritmo, foi efetuada uma análise da cobertura do código. Essa análise possui como propósito verificar se cada requisito foi considerado durante o desenvolvimento.

O trecho de código 3 exibe a consulta ao banco para coletar os parâmetros de entrada do algoritmo. É possível observar que a consulta procura através das tabelas `disciplinas` e `turmas`, procura na tabela `semestres` (que contém as disciplinas dos currículos), procura nas tabelas de avaliações, procura na tabela `historicos` e utiliza-se de um filtro com as disciplinas e turmas já escolhidas pelo aluno, considerando os requisitos RF01, RF02, RF03, RF04 e RF05. A fórmula matemática considera o requisito RF06. Portanto, todos os requisitos funcionais do algoritmo foram considerados.

Como a implementação do algoritmo é um modelo matemático, pode-se

observar que o requisito RNF2 (que determina que o algoritmo deve retornar os mesmos resultados para as mesmas entradas) foi considerado.

A lista a seguir contém os requisitos do sistema, o(s) componente(s) que os consideram, e a rota da API que é utilizada para alimentar a interface.

- **RF07:** O sistema deve permitir que o usuário crie uma grade horária para o próximo período.
 - **Interface:** `grade/+page.svelte`
 - **API:** `/pesquisa/{info,pode cursar,faltacursar,cursadas}`
- **RF08:** O sistema deve permitir que o usuário submeta seu histórico escolar.
 - **Interface:** `HistoricoPopup.svelte`
 - **API:** `/historico`
- **RF09:** O sistema deve permitir que o usuário selecione turmas das disciplinas para compor sua grade horária.
 - **Interface:** Componentes em `components/grade/turma`
 - **API:** `/disciplinas/info`
- **RF10:** O sistema deve permitir que o usuário armazene a sua grade horária finalizada.
 - **Interface:** `JanelaSalvar.svelte` e `GrupoBotoes.svelte`
 - **API:** `/grade`
- **RF11:** O sistema deve permitir que o usuário compartilhe a sua grade horária finalizada.
 - **Interface:** `JanelaSalvar.svelte`
 - **API:** `/grade`
- **RF12:** O sistema deve permitir que o usuário recupere uma grade horária montada a partir de um link compartilhado.
 - **Interface:** `grade/+page.svelte`
 - **API:** `/grade`
- **RF13:** O sistema deve permitir que o usuário inicie uma sessão autenticada utilizando sua matrícula.
 - **Interface:** `CheckLogin.svelte` e `LoginPopup.svelte`
 - **API:** `/login`

- **RF14:** O sistema deve permitir que o usuário finalize uma sessão autenticada.
 - **Interface:** `LoginHeader.svelte`
 - **API:** `/logout`
- **RF15:** O sistema deve permitir que o usuário avalie uma disciplina.
 - **Interface:** `avaliacao/+page.svelte`
 - **API:** `/avaliacoes/`
- **RF16:** O sistema deve permitir que o usuário avalie um professor de uma disciplina.
 - **Interface:** `CampoAvaliacao.svelte`
 - **API:** `/avaliacoes/professor`
- **RF17:** O sistema deve permitir que o usuário altere uma avaliação feita anteriormente.
 - **Interface:** `CampoAvaliacao.svelte`
 - **API:** `/avaliacoes/disciplina`

Os requisitos RF18 e RF19 foram implementados em um serviço diferente, executados periodicamente através da ferramenta `cron` do sistema operacional *Linux*. O serviço é um único script Python. Esse script acessa o microhorário e faz as alterações necessárias no banco de dados.

11.2

Testes da interface

Após verificar que os requisitos foram cobertos durante a fase de desenvolvimento, os requisitos foram testados individualmente. Cada implementação dos requisitos foi testada manualmente com os diferentes tipos de entradas possíveis.

Por exemplo, o requisito RF10 (armazenamento de uma grade) foi testado com diversas grades diferentes, inclusive grades vazias, verificando se a grade era armazenada corretamente no banco de dados. E o requisito RF11 era testado verificando se o código da grade, gerado ao armazená-la no banco de dados, era corretamente fornecida para a interface. Os outros requisitos passaram por testes semelhantes.

Como o código foi desenvolvido de forma a produzir a maioria dos erros em tempo de compilação e a partir de análises semânticas (por utilizar tipos em suas linguagens), não ocorreram muitos erros durante a etapa de testes manuais. Alguns outros poucos problemas simples foram encontrados durante a etapa de entrevistas e foram corrigidos prontamente.

11.3

Teste de performance

Foram efetuados testes de performance para verificar se os requisitos RNF1 e RNF3 foram considerados. A figura 11.1 representa um teste de performance, com 40 usuários simultâneos autenticados solicitando recomendações diferentes a cada 1-5 segundos. É possível observar que o tempo médio de resposta das recomendações foi 20 milissegundos, e não houve nenhuma falha no sistema ou algoritmo. O teste de performance foi desenvolvido usando o framework *Locust* [17], utilizando a linguagem Python.

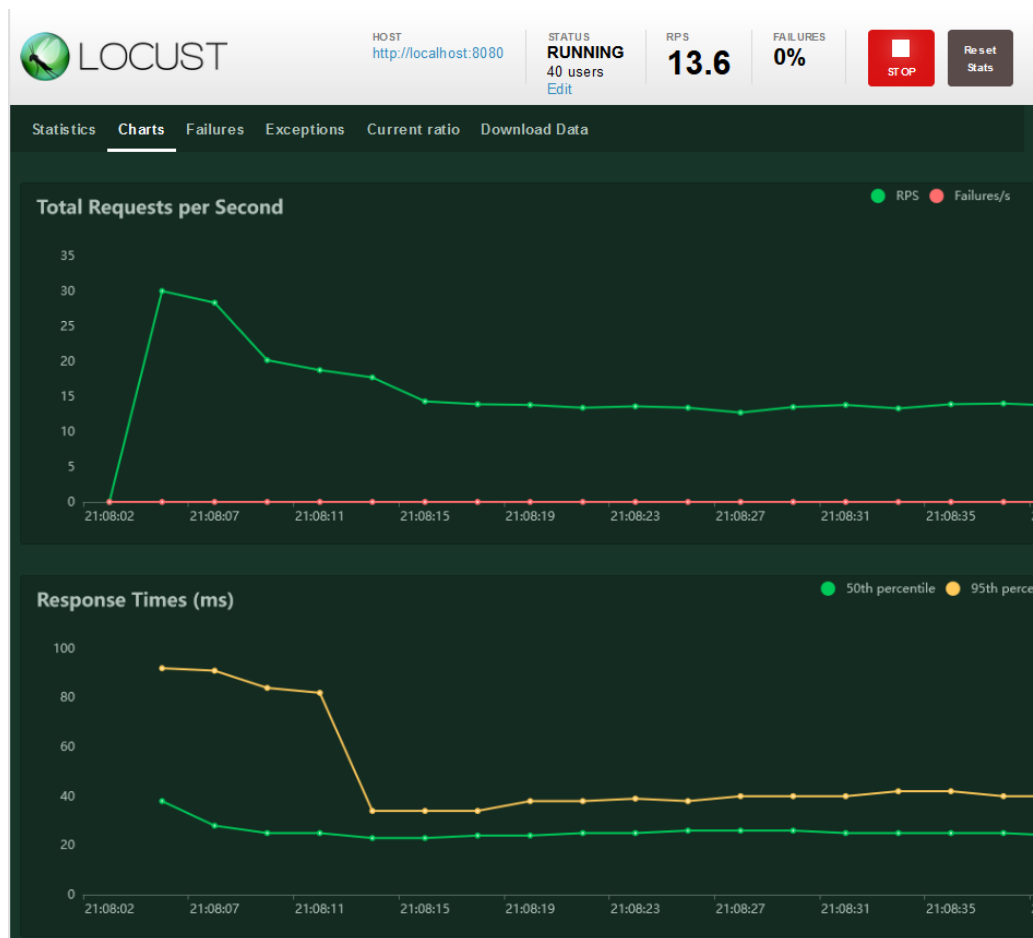


Figura 11.1: Teste de performance da recomendação

12

Uso real

Ao ter-se uma versão com todas os requisitos atendidos, o processo de testes com os usuários foi iniciado. A metodologia da simulação de uso real foi:

1. Apresentar a ideia do sistema e do algoritmo de recomendação;
2. Apresentar todas as possíveis funcionalidades do sistema;
3. Criar uma grade horária como exemplo, utilizando-se de todas as funcionalidades do sistema;
4. Permitir que o usuário faça sua autenticação e crie um grade horária;
5. Ouvir o feedback do usuário.

Os testes foram realizados durante o processo de desenvolvimento e aprimoramento da interface. Por isso, alguns testes foram prejudicados pela qualidade da interface, que possuía todas as funcionalidades previstas, mas não de forma clara.

Foram entrevistados seis alunos do Departamento de Informática, três de engenharia de computação e três de ciência da computação. Cada aluno estava em um período diferente da sua trajetória, variando do sexto semestre do curso até mais de dez semestres.

Os resultados foram satisfatórios. Todos aprovaram a ideia do algoritmo de recomendação, do sistema de auxílio à criação da grade horária e do sistema de avaliação de disciplinas e professores. A seguir estão algumas sugestões fornecidas pelos usuários após testarem o sistema que não foram modeladas como requisitos.

1. **Exibir quantas disciplinas curriculares são trancadas por uma disciplina:** Essa sugestão consiste em exibir a quantidade de disciplinas que uma disciplina bloqueia através dos seus pré-requisitos. Esse número poderia inclusive ser utilizado na fórmula de recomendação afetando P_c . Porém ela não foi implementado devido a complexidade do cálculo, que deve ser implementado de maneira recursiva e de preferência, pré-calculado durante o momento de inserção de novas disciplinas.
2. **Mostrar a quantidade de créditos na grade sendo construída:** Essa sugestão consiste em mostrar a soma dos créditos das disciplinas das turmas selecionadas na grade. Essa sugestão, por sua facilidade de

implementação, foi adicionada pouco tempo depois da entrevista. A quantidade de créditos pode ser visualizada acima do botão de salvar grade horária na interface final.

3. **Exibir quais disciplinas já foram cursadas na lista lateral:** Essa sugestão consiste em diferenciar as disciplinas já cursadas na lista lateral. Essa sugestão foi julgada como importante, e como já havia uma rota na API que disponibiliza quais disciplinas já foram cursadas pelo aluno, esta sugestão foi implementada pouco tempo depois.
4. **Filtrar disciplinas por departamento:** Essa sugestão consiste em permitir filtrar as disciplinas na lista lateral pelos seus departamentos. O sistema atual armazena os departamentos das disciplinas, mas não os utiliza em nenhum lugar. Como essa sugestão envolve várias alterações na interface para permitir essa funcionalidade, essa sugestão será implemenada em um momento futuro.
5. **Exibir disciplinas não recorrentes:** Nem todas as disciplinas eletivas são oferecidas em todos os períodos. Essa sugestão consiste em exibir dos períodos oferecidos da disciplina, e destacá-la caso ela não tenha sido oferecida no último semestre. Como essa sugestão envolve uma reformulação geral das tabelas do banco de dados, das consultas à API e da interface, optou-se por não implementar essa sugestão nesse momento.

Foram feitas outras sugestões que envolviam pequenas alterações na interface que não afetavam o funcionamento do sistema ou algoritmo, que foram implementadas ou não de acordo com o nível de complexidade.

13

Conclusão

A partir da construção e dos testes com os usuários, entendeu-se que objetivo do projeto de auxiliar alunos em seus períodos de matrícula foi alcançado. Ter um sistema para planejar sua matrícula com todas as informações necessárias organizadas em um só lugar, em vez de em três ou quatro serviços diferentes, é inovador para esta universidade. O algoritmo de recomendação cumpre o seu papel em recomendar disciplinas se utilizando de diversas fontes de informação, facilitando o planejamento do usuário aluno da universidade.

Durante todo o processo de entrevistas, modelagem e construção, tanto o sistema como o algoritmo passaram por diversas mudanças. As entrevistas revelaram que os processos de preferência de cada aluno são diferentes, e que mesmo tratando-se de um só departamento, cada aluno tem sua preferência pessoal, de forma que nenhum sistema ou algoritmo pode satisfazer todos os usuários de forma completa e eficaz.

13.1

Estudos futuros

O projeto do sistema possui uma gama de possíveis estudos futuros. Em um primeiro momento, deve-se implementar as funcionalidades descobertas durante o processo de desenvolvimento, como grupo de disciplinas optativas e co-requisitos, e as sugestões complexas provenientes dos usuários na fase de testes. Além disso, foram planejados algumas temas a serem estudados, conforme a lista a seguir:

1. **Disciplinas optativas e co-requisitos:** Conforme explicado nas seções 10.2.2 e 10.2.3, o projeto não considera disciplinas optativas nem co-requisitos de disciplinas. Uma reformulação da arquitetura dos dados é uma oportunidade de estudo para adequar essas características ao sistema.
2. **Exibir quantas disciplinas curriculares são trancadas por uma disciplina:** Essa sugestão foi oferecida por um dos entrevistados durante a fase de análise de uso real. Um estudo focado em desenvolver um algoritmo para determinar a quantidade de disciplinas bloqueadas por uma disciplina curricular é importante para ajudar a determinar o grau de relevância de uma disciplina.

3. **Exibir disciplinas não recorrentes:** Essa sugestão também foi fornecida por um entrevistados durante a fase final. Remodelar o banco de dados de forma a armazenar as disponibilidades das disciplinas conforme os semestres é uma oportunidade de determinar quais disciplinas não são recorrentes, e indicar ao usuário quais disciplinas podem não ser oferecidas no próximo semestre.
4. **Inteligência artificial:** O algoritmo foi construído através de um modelo matemático estático, com pesos fixos. Um estudo focado em inteligência artificial para modelar um novo algoritmo ou uma nova escolha de pesos pode otimizar as recomendações.
5. **Performance da API:** Mesmo satisfazendo os requisitos planejados, a performance da API pode ser melhor otimizada. Melhores consultas, utilização de cache, e visões materializadas são possíveis campos de estudo importantes em um tempo futuro.
6. **Geração automática da grade:** Após validar com um maior grupo de usuários e verificar a robustez das recomendações, poderia haver uma funcionalidade de escolher automaticamente a melhor recomendação de forma automática, gerando uma grade somente com as disciplinas recomendadas.
7. **Estudar outros pesos:** Apesar das entrevistas revelarem cinco categorias de uma escolha de uma disciplina, essas podem não ser as melhores para modelar uma recomendação. Um estudo futuro pretende efetuar mais entrevistas, coletar mais informações e possivelmente adicionar ou remover pesos do algoritmo, de forma a otimizar a recomendação das disciplinas.

13.2

Código completo

Todo os códigos desenvolvidos para o projeto estão disponíveis publicamente em repositórios do Github:

- github.com/Leinadium/puc-disciplinas: Código do sistema e do algoritmo e suas respectivas documentações;
- github.com/Leinadium/microhorario-dl: Código da biblioteca para interação com o microhorario;
- github.com/Leinadium/puc-tcc-disciplinas: Código para a criação e documentação do relatório final.

Referências bibliográficas

- 1 PUC-RIO, S. A. U. *Nova matrícula de graduação PUC-Rio*. 2018. Disponível em: <https://www.cbctc.puc-rio.br/Publicacao/Paginas/Files/20181/Apresenta%C3%A7%C3%A3oNovaMatr%C3%ADculaPUC-Rio_09jul18.pdf>. Acesso em: 25 abr. 2023.
- 2 NG, Y.-K.; LINN, J. Crsrecs: A personalized course recommendation system for college students. In: *2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA)*. [S.l.: s.n.], 2017. p. 1–6.
- 3 RANI, L. P. J. et al. Course recommendation for students using machine learning. In: *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*. [S.l.: s.n.], 2020. p. 381–384.
- 4 NGUYEN, V. A. et al. A course recommendation model for students based on learning outcome. In: *2021 Education and Information Technologies*. [S.l.: s.n.], 2021. (26), p. 5389–5415.
- 5 ADAK, M. F.; YUMUSAK, N.; TASKIN, H. An elective course suggestion system developed in computer engineering department using fuzzy logic. In: *2016 International Conference on Industrial Informatics and Computer Systems (CIICS)*. [S.l.: s.n.], 2016. p. 1–5.
- 6 XU, J.; XING, T.; SCHAAR, M. van der. Personalized course sequence recommendations. *IEEE Transactions on Signal Processing*, v. 64, n. 20, p. 5340–5352, 2016.
- 7 ADAK, M. F.; ERCAN, S. Support vector machine and decision tree-based elective course suggestion system: A case study. In: *2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*. [S.l.: s.n.], 2021. p. 552–556.
- 8 CHEN, P. P.-S. The entity-relationship model—toward a unified view of data. *ACM Trans. Database Syst.*, Association for Computing Machinery, New York, NY, USA, v. 1, n. 1, p. 9–36, mar 1976. ISSN 0362-5915. Disponível em: <<https://doi.org/10.1145/320434.320440>>. Acesso em: 21 set. 2023.
- 9 The PostgreSQL Global Development Group. *PostgreSQL: The World's Most Advanced Open Source Relational Database*. 2023. Disponível em: <<https://www.postgresql.org/>>. Acesso em: 02 nov. 2023.
- 10 Django Software Foundation. *The web framework for perfectionists with deadlines / Django*. 2023. Disponível em: <<https://www.djangoproject.com/>>. Acesso em: 02 nov. 2023.
- 11 Pallets. *Welcome to Flask*. 2023. Disponível em: <<https://flask.palletsprojects.com/>>. Acesso em: 02 nov. 2023.

- 12 Gin Team. *Gin Web Framework*. 2023. Disponível em: <<https://gin-gonic.com>>. Acesso em: 02 nov. 2023.
- 13 Sergio Benitez. *Rocket - Simple, Fast, Type-Safe Web Framework for Rust*. 2023. Disponível em: <<https://www.rocket.rs>>. Acesso em: 03 nov. 2023.
- 14 The Rust Foundation. *Rust Programming Language*. 2023. Disponível em: <<https://www.rust-lang.org>>. Acesso em: 03 nov. 2023.
- 15 Google. *The Go Programming Language*. 2023. Disponível em: <<https://go.dev>>. Acesso em: 02 nov. 2023.
- 16 Svelte Contributors. *SvelteKit - Web development, streamlined*. 2023. Disponível em: <<https://kit.svelte.dev>>. Acesso em: 03 nov. 2023.
- 17 C. Bystrom, J. Heyman, J. Hamrém, H. Heyman e L. Holmberg. *Locust - A modern load testing framework*. 2023. Disponível em: <<https://locust.io>>. Acesso em: 04 nov. 2023.