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



Daniel Schreiber Guimarães

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

Relatório de Projeto Final

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 titulo de Bacharel em Engenharia da Computação.

Orientador: Prof. Marcos Vianna Villas

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.

23 f: il. color.; 30 cm

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

Inclui bibliografia

 Informática – Teses. 2. Recomendação. 3. Disciplinas.
 Villas, Marcos Vianna. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

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. 23p. Projeto Final – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

This abstract will be filled.

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. 23p. Projeto Final – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Esse resumo ainda será preenchido.

Palavras-chave

Recomendação; Disciplinas.

Sumário

1	Introdução	9
2	Situação Atual	10
2.1	Serviços disponibilizados	10
2.2	Processo de matrícula	11
2.3	Estudos relacionados	11
3	Proposta	13
4	Plano de Ação	14
4.1	Cronograma original	14
4.2	Cronograma revisado	15
5	Validação do problema	16
5.1	Método das entrevistas	16
5.2	Perguntas realizadas	16
5.3	Resultados das entrevistas	17
6	Pré requisitos	22
6.1	-	22
7	Referências bibliográficas	23

Lista de figuras

O	Interface do <i>microhorario</i> Interface do simulador	10 12
•	Cronograma original do projeto Cronograma atualizado do projeto	14 15
Figura 5.1 dos atuais.	Distribuição da quantidade de entrevistados pelos seus perío-	18

Lista de tabelas

Tabela 5.1	Graus dos critérios de escolhas de disciplinas escolhidos pelos	
entrevistados		19
Tabela 5.2	Relação de peso para grau dos critérios de escolhas	20
Tabela 5.3	Pesos dos critérios de escolhas de disciplinas escolhidos pelos	
entrevistados.		20

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 pelo aluno.

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.



Figura 2.1: Interface do *microhorario*

Para consultar a grade recomendada e os pré-requisitos, o aluno precisa acessar um documento² presente na página do departamento ou acessar o currículo disponibilizado na página da universidade. 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,

¹https://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.

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. As datas possíveis abrangem um período de cinco dias. Durante o período de matrícula, o microhorario 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.

A universidade disponibiliza um simulador de matrícula por um curto período 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 simulador apresenta os dados das disciplinas atualizados conforme o microhorario, 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 for-



Figura 2.2: Interface do simulador

necidas, 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 assemelham ao 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.

Proposta

O sistema proposto nesse trabalho é uma interface de planejamento de matrícula semelhante ao simulador de matrícula. O sistema disponibiliza as disciplinas do próximo período conforme os dados mais atuais do microhorario. Para que haja uma personalização na recomendação das disciplinas, o sistema permitirá 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 receberá as disciplinas oferecidas no próximo período e seus pré-requisitos, o modelo de grade recomendada pelo departamento e o histórico do aluno e responderá com disciplinas selecionadas pelo algoritmo. O algoritmo também poderá ser personalizado com informações de preferência do usuário que serão definidas ao longo da etapa de validação do problema.

Também será validada a proposta de armazenar o histórico fornecido pelo aluno sem os 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 será 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 permite restrigir o escopo do projeto afim de tentar obter melhores resultados.

4

Plano de Ação

Para modelar o sistema, em um primeiro momento será feita uma validação do problema. Alunos do departamento de informática da universidade serão 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 é 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 passará por uma etapa de pesquisa. Serão 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 serão analisadas as fontes de informações disponibilizadas pela universidade e como cada fonte pode ser integrada no sistema. Por exemplo, já existe uma biblioteca¹ Python para acessar e processar as disciplinas atualmente no microhorario.

Após a etapa de pesquisa, o algoritmo começará a ser desenvolvido. Depois de obter um algoritmo minimamente viável, será desenvolvida a interface de planejamento para mostrar o funcionamento do algoritmo. Tanto o algoritmo com a interface serão desenvolvidas utilizando um modelo de desenvolvimento incremental, para que o sistema passe por etapas de desenvolvimento, teste e validação com os alunos.

4.1 Cronograma original

O cronograma inicial pode ser visualizado na figura 4.1.

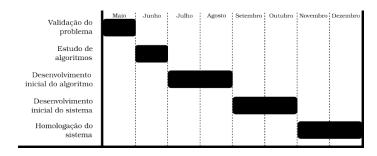


Figura 4.1: Cronograma original do projeto

¹Disponível em: https://pypi.org/project/microhorario-dl/

4.2 Cronograma revisado

O plano de ação original foi levemente modificado. A validação do problema durou um pouco mais do que planejado. As etapas relacionadas ao algoritmo foram agrupadas no desenvolvimento do sistema, e foi criada uma nova etapa de elaboração de requisitos. O cronograma revisado pode ser visualizado na figura 4.2.

Figura 4.2: 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 as suas preferências ao realizar as suas matrículas. As opiniões são 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. Essa média pode ser usada pelo algoritmo para saber quantas

disciplinas o estudante ainda pretende fazer naquele semestre e recomendar de acordo.

- 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, pois a ferramenta com o algoritmo estaria disponibilizada em um sistema cujo obtivo é planejar a sua matrícula.
- 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, por exemplo.

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

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. 5 dos 6 alunos entrevistados de engenharia de computação já haviam cursado mais da metade do curso (já cursou 5 dos 10 semestres recomendados), e 3 dos 5 alunos entrevistados de ciência de computação já haviam cursado mais da metade do curso (já cursou 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 proximo 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 entrevisto 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.".

Como pode ser observado na resposta fornecida acima, as respostas podem variar bastante de aluno para aluno. Por isso, 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 com 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
Entrevistado 1	3	-	-	1	2
$Entrevistado\ 2$	1	-	-	3	2
Entrevistado 3	2	3	-	-	1
Entrevistado~4	3	1	-	2	-
$Entrevistado\ 5$	2	-	-	1	3
$Entrevistado\ 6$	1	-	-	_	_
Entrevistado 7	1	-	-	-	-
$Entrevistado\ 8$	2	3	1	4	_
Entrevistado~9	3	2	-	_	1
$Entrevistado\ 10$	1	_	3	2	_
Entrevistado 11	-	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 de acordo com a tabela 5.2.

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.

Grau	1	2	3	4	5	-
Peso	5	4	3	2	1	0

Tabela 5.2:	Relação o	de peso	para gran	dos	critérios	de	escolhas
100010 0.2.	i coração c	TO PODO	para Sraa	aob	CITOCITOD	ac	Cocomina

	Conteúdo	Professor	Avaliação	Horário	Opinião
Entrevistado 1	3	0	0	5	4
$Entrevistado\ 2$	5	0	0	3	4
Entrevistado 3	4	3	0	0	5
Entrevistado~4	3	5	0	4	0
$Entrevistado\ 5$	4	0	0	5	3
$Entrevistado\ 6$	5	0	0	0	0
Entrevistado 7	5	0	0	0	0
Entrevistado~8	4	3	5	2	0
Entrevistado 9	3	4	0	0	5
Entrevistado~10	5	0	3	4	0
Entrevistado 11	0	4	0	5	0
Soma	41	19	8	28	21

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

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 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 de facilidade e relevância. Dois entrevistados disseram que inicialmente escolhem 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

21

 $me\ preocupar.\ Eu\ encaro\ as\ [disciplinas]\ eletivas\ como\ um\ escape\ para\ poder$ respirar."

6 Pré requisitos

6.1 Do algoritmo

Os pré-requisitos do algoritmo de recomendação de disciplinas precisa satisfazer os critérios 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. Portanto, os requisitos abaixo planejam abranger

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.I.: 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.I.: 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.I.: s.n.], 2021. p. 552–556.