

Um Modelo para Recomendação de Cursos de Especialização Baseado no Perfil Profissional do Candidato

Antônio Eduardo Rodrigues de Souza¹, Sandra Maria Dotto Stump¹

¹Programa de Pós-graduação em Engenharia Elétrica
Universidade Presbiteriana Mackenzie (UPM) – São Paulo – SP - Brasil
aersouza@gmail.com, sstump@mackenzie.br

Abstract: *This paper proposes to study professional profiles of candidates on choosing a specialization course. A methodology based on the process Knowledge Discovery in Databases (KDD) is applied, and proposed a course recommendation model, using a data filtering technique based on collaboration with the use of data mining to extract knowledge from the database. On this project, it is expected to be detected the courses which best suits each candidate profile, giving support to academic institution to satisfy candidates needs and reduce the number of dropouts or changes.*

Resumo: Este artigo propõe estudar os perfis profissionais de candidatos na escolha de um curso de especialização. Aplicar-se-á uma metodologia baseada no processo Knowledge Discovery in Databases (KDD), e proposto um modelo de recomendação, utilizando-se uma técnica de filtragem de dados baseada na colaboração, além do uso da mineração de dados na extração de conhecimento da base de dados. Com este projeto, espera-se serem detectados os cursos mais adequados ao perfil de cada candidato, reduzindo-se o número de desistências ou mudanças.

1. Introdução

As novas necessidades da sociedade atual mostram que o percentual de cursos oferecidos aumentou, visando atender às demandas do mercado. Mostra também que se busca qualidade nos cursos ofertados, que têm como objetivo principal preparar profissionais mais capacitados para o mercado de trabalho.

O presente trabalho busca, a partir de informações históricas dos candidatos de cursos de espe-



Secretaria de Educação
Profissional e Tecnológica



Ministério
da Educação





cialização de uma universidade particular, situada na cidade de São Paulo, identificar e analisar perfis de comportamento. A pesquisa estará delimitada aos candidatos que procuraram cursos de especialização, durante vários semestres consecutivos, nos últimos anos.

A universidade realiza semestralmente a seleção de candidatos a cursos de especialização *latu sensu*, numa quantidade de, aproximadamente, três mil candidatos inscritos a cada período. O processo seletivo é composto por um formulário eletrônico, disponibilizado no portal da universidade, contendo questões que, além das informações pessoais, permite aprofundamento de conhecimento da formação, objetivo da procura do curso, expectativas em relação ao conteúdo e as razões da procura, como por exemplo, atualização profissional, promoção, etc. As respostas de cada candidato são gravadas em base de dados. De posse destas informações, o coordenador de cada curso verifica as informações armazenadas de cada candidato e avalia se o curso escolhido é realmente o mais adequado ao perfil informado. Somente no momento da divulgação dos resultados dos candidatos aprovados é que se tem conhecimento do curso onde estará matriculado, uma vez que o formulário eletrônico permite ao candidato selecionar três opções de cursos.

Objetiva-se desenvolver um modelo baseado em perfis profissionais de candidatos, utilizando-se a inteligência artificial, técnicas de filtragem e de mineração de dados, para personalizar ofertas de cursos de especialização.

O estudo torna-se relevante à medida que possa oferecer aos candidatos, que procuram os cursos de especialização, opções de cursos adequadas às suas expectativas.

Este artigo está dividido em 2 seções: a primeira expõe a problemática que se pretende estudar e sua justificativa. A seção 2 apresenta as etapas para se chegar ao modelo proposto.

2. O modelo proposto

Será realizada, inicialmente, uma pesquisa documental, por meio de levantamento bibliográfico, onde se pretende conhecer o estado da arte dos sistemas de recomendação, avaliar as técnicas de filtragem e mineração de dados disponíveis, e adaptá-las ao modelo proposto.

O trabalho será desenvolvido a partir da coleta de dados de uma pesquisa exploratória, com abordagem qualitativa, entre os que foram candidatos a cursos de especialização em uma universidade





particular, na cidade de São Paulo, entre os anos de 2006 e 2011.

Para a descoberta de conhecimento foram estudados os processos CRISP-DM [Chapman *et al.* 2000] e o KDD [Fayyad *et al.* 1996] (Figura 1). Optou-se pela utilização do processo KDD, pela melhor aderência às necessidades do modelo proposto.

2.1. Etapas do Processo KDD

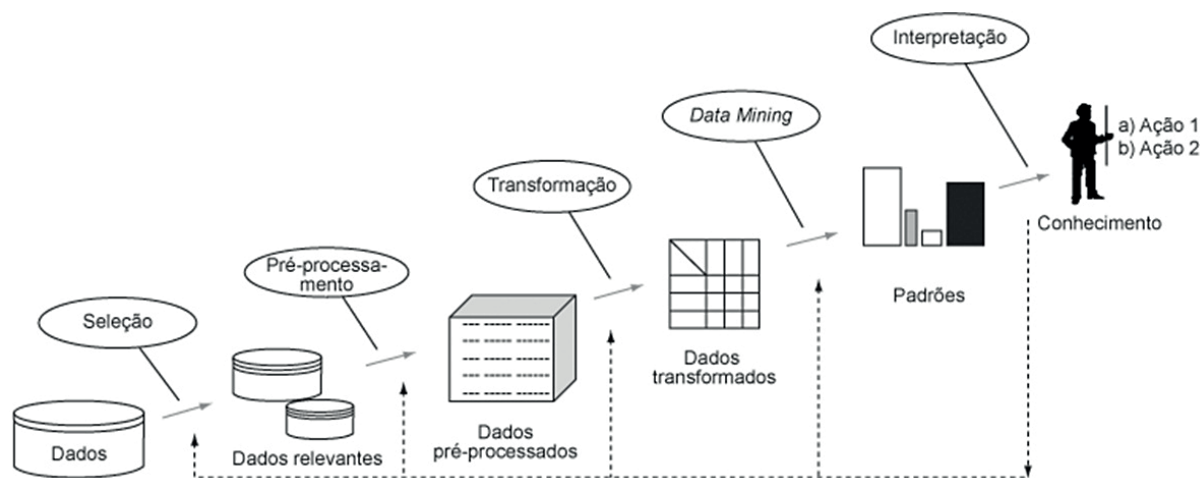


Figura 1: Etapas do processo KDD [FAYYAD et al. 1996]

Primeiramente será realizado o entendimento e a definição de um domínio. Em seguida, serão selecionados, dentro do domínio estabelecido, os dados nos quais a descoberta será realizada. Dentre os dados adquiridos, serão consideradas, para análise, as informações dos perfis profissionais dos candidatos, constantes no formulário de inscrição. O perfil profissional é um conjunto de características que precisam ser encontradas em um candidato para que este possa ocupar um determinado cargo, podendo ser dividido em conhecimentos técnicos e perfil comportamental (Tabela 1):



Tabela 1: Fonte [Santos et al. 2004]

Conhecimentos Técnicos	Perfil Comportamental
Escolaridade	Comunicação
Área de formação	Relacionamento interpessoal
Idiomas	Bom senso
Experiência profissional	Postura
Informática	Ética

Em razão da subjetividade e dificuldade encontradas na mensuração dos valores que expressam o perfil comportamental, será considerado neste trabalho apenas variáveis que representam o conhecimento técnico (Tabela 2).

Tabela 2: Variáveis (atributos) utilizadas no estudo

Variáveis	Descrição	Exemplo
Graduação	Nome do curso de graduação	Administração de Empresas, Arquitetura e Urbanismo, Ciência da Computação
Habilitação	Titulação do curso	Bacharelado, licenciatura, tecnólogo, médico
Pós-graduação	Nome do curso de pós-graduação	Administração de Empresas, Arquitetura e Urbanismo, Ciência da Computação
Atividade profissional	Nome da atividade que exerce	Analista de Sistemas, Arquiteto, Controller, Jornalista, Secretária
Cargo que ocupa	Nome do cargo que ocupa	Assessor, Analista, Auxiliar, Supervisor, Coordenador, Gerente
Idioma	Conhecimento no idioma inglês	Básico, intermediário, avançado, fluente





Como forma de preparação para a etapa seguinte, estes dados devem ser limpos e transformados. Essa limpeza inclui a remoção de ruídos, que são dados com erros ou valores discrepantes, a adequação de valores que estejam fora de contexto, a inclusão de valores ausentes, a seleção e o resumo das variáveis a serem utilizadas. Valores ausentes, ou seja, não constantes nas informações dos registros, serão inseridos utilizando-se uma constante global, a média de cada atributo, ou a média de todas as variáveis de mesma classe. No caso de atributos nominais será utilizado o subconjunto dominante, enquanto que no caso de atributos não nominais será utilizada a moda de cada classe, que é o valor que ocorre com maior frequência em um conjunto de dados.

Ainda nesta etapa, serão definidas as técnicas e algoritmos de mineração de dados que serão utilizados, selecionados do domínio que deverão ser então transformados de acordo com as características técnicas dos algoritmos. Com isso, serão submetidos à etapa de mineração propriamente dita. Pretende-se utilizar regras de associação para a classificação dos dados, por apresentarem melhor precisão na recomendação de cursos [Bendakir e Aimeur 2000].

Com a base de dados preparada, será realizada a etapa de mineração e aprendizagem de máquina, utilizando-se a ferramenta WEKA (Waikato Environment for Knowledge Analysis), desenvolvida pela Universidade de Waikato, na Nova Zelândia [Hall et al. 2009]. Esta ferramenta foi selecionada por ser de domínio público e trabalhar com diversas técnicas de mineração de dados, tais como regras de associação, agrupamento, classificação, e diversos algoritmos.

A última fase compreende o desenvolvimento de um relatório (Tabela 3) para exibição de resultados, permitindo a universidade identificar quais perfis profissionais melhor se adéquam a determinados cursos, considerando, inclusive, possíveis remanejamentos.



Tabela 3: Protótipo do relatório de saída de dados

Recomendação de Cursos de Especialização baseado no Perfil Profissional do Candidato			
Graduação / Habilitação	Atividade profissional / Cargo	Inglês	Curso desejado
Ciência da Computação	Analista de Sistemas	Fluente	Gerenciamento de Projetos
Bacharelado	Gerente		
Curso recomendado: Gerente de Projeto			
Ciência da Computação	Analista de Sistemas	Básico	Gerenciamento de Projetos
Bacharelado	Analista		
Curso recomendado: Projeto e Desenvolvimento de Sistemas			
Administração de Empresas	Analista de Projetos	Avançado	Gerenciamento de Projetos
Bachelor	Supervisor		
Curso recomendado: Governança de TI			

Como resultado gerado, o modelo adquirido será analisado pelo coordenador pedagógico de cada curso, para melhorar a compreensão do conhecimento descoberto pelo algoritmo minerador. Caso os resultados da mineração não sejam satisfatórios, várias etapas do processo poderão ser realizadas novamente para que se consiga agregar informações que contribuam no processo de descoberta de conhecimento.

Referências

Bendakir, N. and Aimeur, E. (2006) "Using Association Rules for Course Recommendation".

Chapman, P. et al. (2000) "CRISP-DM 1.0 - Step-by-step Data Mining Guide".

Fayyad, U. et al. (1996). From Data Mining to Knowledge Discovery: An Overview. In *Advances in Knowledge Discovery and Data Mining*, [S.I.] AAAI Press.



Santos, D. *et al.* (2004) "Redes Neurais Artificiais Aplicadas em Análise de Perfis Profissionais".

Hall, M. *et al.* (2009). The WEKA Data Mining Software: An Update, SIGKDD Explorations, Volume 11, Issue 1.

