

Seminário Internacional de Educação Superior 2014

Formação e Conhecimento

Anais Eletrônicos



ANÁLISE DO PERFIL DISCENTE DA UNIVERSIDADE DE SOROCABA, MEDIANTE A TÉCNICA DE MINERAÇÃO DE DADOS

Rafael Fernando de Moraes Moreno*

RESUMO

No cenário complexo que constitui a universidade nos dias atuais, impulsionada e pressionada pela lógica do mercado, cabe a perguntar: quem está frequentando a universidade? Existe alguma tendência estatística que define o perfil do aluno mediante a opção do curso escolhido? O objetivo desta pesquisa foi realizar um estudo de caso na Universidade de Sorocaba identificando possíveis padrões e classificações no perfil dos discentes que expliquem a sua escolha por determinado curso. Para este feito foi realizado uma análise empírica no banco de dados da Universidade fazendo uso de um software de mineração de dados e utilizando um processo denominado KDD (Knowledge Discovery in Databases), que tem a finalidade de descobrir conhecimento em grande volume de dados. Com esta técnica e ferramenta de mineração de dados foram analisados os dados socioeconômicos e do questionário informativo dos estudantes ingressantes entre 2006 a 2011, extraindo possíveis classificações que expliquem o perfil do estudante que busca determinado curso.

Palavras-chave: Perfil discente. Universidade de Sorocaba. Mineração de dados.

1 INTRODUÇÃO

Hoje ocorre uma forte presença do mercado operando no âmbito universitário e modificando seus valores primordiais. A universidade é produto de um acontecimento histórico que vem desde a baixa Idade Média (Bolonha, 1088), passando por um longo processo de reconstrução social (BERNHEIM, 2010).

Uma Universidade deve exercer o ensino, a pesquisa e a extensão. Cada um destes elementos estruturais emergiu ao longo dos séculos em diferentes modelos universitários. Pode-se destacar alguns: Universidade de Bolonha, Universidade de Paris, e mais tarde, muitas outras, como a Universidade Napoleônica, a Universidade de Berlim (Humboldt) e as Universidades Estadunidenses. Todas exerceram papel fundamental na construção da Universidade dos dias atuais.

* Analista de sistemas - Universidade de Sorocaba. Mestre em Educação Superior. E-mail: rafael.moreno@prof.uniso.br

Seminário Internacional de Educação Superior 2014

Formação e Conhecimento

Anais Eletrônicos



A Educação Superior vem passando por grandes transformações de seus valores e perdendo sua força/significado pela opressão do cenário neoliberal que constitui a globalização econômica.

Nesse cenário complexo, da universidade dos dias atuais, cabe a pergunta: quem está frequentando a universidade? Existe alguma tendência estatística que define o perfil do aluno mediante a opção do curso escolhido?

Com esses questionamentos, tomou-se a Universidade de Sorocaba como um estudo de caso com objetivo de entender e conhecer melhor o público que está inserido na educação superior e assim analisar e delinear possíveis regras existentes no perfil do aluno sobre o curso escolhido.

Para o desenvolvimento desta pesquisa, trabalhou-se com o banco de dados da Universidade de Sorocaba, relativos aos ingressantes entre 2006 a 2011, aplicando a tecnologia denominada mineração de dados (*data mining*) para buscar possíveis regras do perfil discente quando este opta por um curso.

A mineração de dados efetua a junção da estatística com a inteligência computacional, possibilitando a descoberta de comportamentos ocultos em grandes bancos de dados e, a partir daí, a inferência de regras para previsão de comportamento futuro e orientação na tomada de decisões (NAMEN, 2013 apud Hirji, 2001).

Turban, Rainer e Potter (2005) apresentam um exemplo clássico da utilização da mineração de dados na rede Wall Mart. Neste exemplo foi identificado que a compra de fraldas estava tendencialmente ligada à compra de cervejas. Chegou-se à conclusão que se torna impossível a olho nu verificar uma tendência quando há um grande volume de dados.

A tecnologia de mineração de dados foi originalmente utilizada em grandes empresas para tomada de decisão, mas ela serve perfeitamente para a pesquisa em questão, pois, com ela é possível analisar grande volume de dados e mediante a utilização de algoritmos matemáticos garimpar comportamentos que estavam invisíveis ao olhar humano.

Este estudo apresenta grande importância para melhor compreender o cenário da região de Sorocaba, decifrando quem frequenta a universidade, vista do ponto da Universidade de Sorocaba e se existe alguma tendência no perfil do discente quando este opta pelo curso.

Para realizar esse feito, partimos do pressuposto que, analisando três hipóteses, será possível alcançar o objetivo da pesquisa. A primeira hipótese é que exista uma regra na escolha do curso de acordo com os dados



socioeconômicos/demográficos, ou seja, mediante esses dados do estudante será possível cruzá-los e assim extrair possíveis regras que definam o perfil discente sobre um determinado curso.

A segunda hipótese é que exista uma regra nos dados do processo seletivo, ou seja, de acordo com o questionário informativo que o candidato preenche no dia do vestibular e com outras variáveis como número de acertos, pontuação e outros dados, será possível traçar a regra que define o perfil discente com relação a um determinado curso.

A terceira hipótese diz respeito à combinação dos dados socioeconômicos/demográficos e do processo seletivo na qual será possível cruzar as informações e chegar a uma regra que defina o perfil discente de acordo com o curso escolhido.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Globalização, neoliberalismo e Universidade

Quando tratamos da Universidade e da Educação de um modo geral, seja ela básica ou superior, não podemos deixar de citar um fenômeno que vem ocorrendo e que trouxe consigo mudanças que causaram grande impacto na educação. Estamos falando da globalização.

Se quisermos entender e decifrar o perfil do candidato que está inserido na Universidade, seja na Universidade de Sorocaba ou em qualquer outra instituição, primeiro precisamos compreender os processos mais amplos, políticos e econômicos que operam e exercem influência sobre a Educação e quais são os efeitos que estes produzem.

Com o cenário competitivo instaurado pela globalização, a diluição das fronteiras territoriais e a supervalorização da economia como motor da sociedade, a educação, em especial a educação superior, é forçada a suprir e alimentar o mercado com mão-de-obra capacitada para produzir e consumir, fomentando assim o fortalecimento da economia e a expansão do mercado.

Vale lembrar que a base da economia de mercado é a oferta e demanda. Estando essas duas bases bem alinhadas e proporcionais, a economia tende a ir bem. Trazendo essa lógica para o ponto de vista educacional, uma pessoa analfabeta não consome e uma pessoa sem qualificação mínima não produz.

Seminário Internacional de Educação Superior 2014

Formação e Conhecimento

Anais Eletrônicos



Assim, podemos perceber um alinhamento das políticas impostas pelo Banco Mundial com essa lógica de mercado. Uma destas políticas é o fortalecimento de investimentos em educação básica, capacitando a sociedade minimamente para consumir (demanda). Outra política imposta pelo Banco Mundial resulta no aumento das privatizações em educação superior formando mão-de-obra qualificada para produzir o que podemos chamar de oferta.

Conforme cita Mello (2012, p. 232), de “apenas 2% dos empréstimos do banco mundial para educação, na década de 90 passamos para 22,1% do total de recursos emprestados pelo Banco Mundial ao Brasil no final do período de 1991-1994”.

Foi, sobretudo na década de 90 que o Banco Mundial impôs fortemente ao Brasil o seu modelo, fazendo com que o governo priorizasse os investimentos em educação básica e aumentasse a privatização na educação superior.

Essas trouxeram um acréscimo ao número de pessoas cursando a educação superior. Mas o que mais inquieta e se torna questionável é a forma como isso está ocorrendo, sem entrar no mérito da qualidade desses cursos, pois aí teríamos que questionar de qual qualidade se está tratando. Estamos hoje formando cidadãos-profissionais ou somente profissionais para o mercado de trabalho?

Não podemos nos esquecer de que educar não é simplesmente ensinar uma profissão aos alunos. Educar é mais amplo do que ensinar, conforme escreve Barata-Moura (2003, p. 31), “educar não é domesticar, nem formatar; educar é eduzir: desenvolver e formar a nossa aptidão de senhorio e de orientação do viver”.

Com o aumento de privatizações na educação superior, investidores internos e externos brigam pela fatia do bolo, que se demonstrou ser muito atrativa e lucrativa. Conforme relata Cunha (2004), de acordo com o censo de 2003, houve um aumento de 209% dos alunos que cursam uma instituição superior, mas o número de universidades públicas ficou estagnado, ou seja, tivemos o aumento de instituições privadas e a criação dos centros universitários.

Podemos observar esse aquecimento no “mercado da educação” com a seguinte passagem de Marques (2013, p. 72) “[...] A matrícula na educação superior [...] de 3.036.113 matriculados em 2001, alcança 6.379.299 em 2010”. Subdividindo esse número em público ou privado, Marques (2013, p. 72) detalha que “[...] as públicas aumentaram de 944.584 para 1.643.298, uma taxa de 42,5%, já as privadas de 2.091.529 para 4.736.001, uma taxa de 55,8%”.

Seminário Internacional de Educação Superior 2014

Formação e Conhecimento

Anais Eletrônicos



O surgimento da sociedade do conhecimento vem agravando mais a situação, pois, segundo Dias Sobrinho (2005, p. 110) a chamada sociedade do conhecimento, que é fundada na informação e comunicação, trouxe consigo mais assimetrias sociais. O conhecimento tornou-se motor da economia. Hoje os países ricos e desenvolvidos são aqueles que detêm os meios de comunicação/informação gerando mais conhecimento, detendo as pesquisas de ponta e conseqüentemente obtendo riqueza.

Os avanços tecnológicos e informacionais que movem a economia nos dias atuais fizeram crescer o número de pessoas excluídas. De acordo com Dias Sobrinho (2005, p. 111), cerca de 4,5 bilhões não possuem condições necessárias de acesso às novas tecnologias de comunicação e informação e conseqüentemente a estas serão negados empregos, condições mínimas de vida digna e esperança de futuro.

Uma frase que representa o desafio atual é o que diz Dias Sobrinho (2005, p. 112):

Um dos maiores desafios do mundo de hoje é dar um sentido mais humano ao progresso, isto é, fazer com que os benefícios do desenvolvimento tecnológico e global cheguem ao maior número possível de pessoas em todos os cantos do planeta.

Temos duas forças interagindo na educação: a mercantilização da educação como algo que se possa comprar, equiparando-a com um simples produto/serviço, e a formação de mão-de-obra para o mercado, minimamente qualificada pra executar aquilo que é necessário do ponto de vista empresarial e com necessidades de consumo desenfreado.

Goergen (2013, p. 48) constata que a educação hoje é vista como investimento individual, visando retorno econômico privado na forma de emprego/salário e retorno estatal em termos de impostos e crescimento econômico. Ou seja, a educação é vista como um serviço que se compra para proporcionar ao indivíduo melhor poderio de competitividade visando sua própria ascensão, não se preocupando com os outros ao redor e nem com a sociedade em que está inserido.

De modo geral, apenas ensinar (adestrar) as pessoas para exercerem uma profissão não é suficiente. Goergen (2013, p. 46) expõe a necessidade da formação e não apenas do ensino. Formação é algo mais amplo que ensino, pois inclui dimensões epistêmicas, políticas e éticas dos alunos.

Diante desse quadro da educação, seja ela superior ou básica, não podemos deixar de acreditar que esta cumpra seu papel de transformar e formar pessoas, que serão capazes de pensar, agir e propor mudanças em nossa sociedade.



Frente a todas essas dificuldades que encontramos na educação, particularmente falando da educação superior, emerge a pergunta: Qual é a missão da educação superior? Vemos hoje um forte apelo para a formação de mão-de-obra, mas será que sua missão é apenas capacitar para o mercado?

Um trecho que Goergen (2013, p. 49) cita traz a definição de qual seria, segundo a UNESCO, a missão da educação superior, sintetizada nas conclusões da Conferência Mundial do Ensino Superior em Paris em 1998:

[...] A missão do ensino superior consiste em contribuir com o desenvolvimento e a melhoria do conjunto da sociedade, ou seja, educar e formar graduados altamente qualificados, capazes de responder às necessidades de todos os aspectos da atividade humana; promover, gerar e difundir conhecimentos através da pesquisa; interpretar, preservar e promover as culturas no contexto do pluralismo e da diversidade cultural; oferecer possibilidades de aprendizagem ao longo de toda a vida; contribuir com o desenvolvimento e melhoria da educação em todos os níveis; proteger e promover a sociedade civil, formando os jovens de acordo com os valores nos quais se baseia a sociedade civil democrática e que proporcionem perspectivas críticas e independentes no debate sobre as opções estratégicas e no fortalecimento das perspectivas humanísticas.

2.2 Descoberta de conhecimento em banco de dados

Nesta pesquisa foi utilizado um processo denominado KDD (Knowledge Discovery in Databases), que em português quer dizer Descoberta de Conhecimento em Banco de Dados, este processo compreende a transformação de dados em conhecimento, para que isso seja possível o KDD apresenta as seguintes etapas:

- **Seleção** - A primeira etapa do processo da KDD é a seleção, é nesta etapa que definiremos a segmentação dos dados necessários para a análise de acordo com critérios pré-definidos. No caso desta pesquisa, foi neste passo que filtramos somente os dados dos alunos que ingressaram no ano de 2006 a 2011.
- **Pré-Processamento** - A segunda etapa do KDD diz respeito à limpeza dos dados selecionados na etapa anterior, com relação à pesquisa, foi aplicada uma série de procedimentos para limpar os dados inconsistentes.
- **Transformação** – Dependendo de qual informação precisamos trabalhar, é necessário transformar alguns dados, por exemplo, foi necessário saber a faixa etária dos alunos, dado este que foi transformado da data de nascimento para a faixa etária, ou seja, é nesta etapa que transformamos os dados em um padrão que necessitamos para trabalhar.



- **Data-Mining** – É nesta etapa que se faz o uso de uma ferramenta (software) para aplicar as técnicas de mineração (algoritmos) nos dados já definidos das etapas anteriores.
- **Interpretação** – Esta é a última etapa do KDD, é nesta fase que interpretamos o resultado extraído da mineração, caso o resultado seja irrelevante, será necessário refazer todas as etapas anteriores.

2.2.1 Técnicas de mineração de dados

Existem vários tipos de conhecimento que podem ser extraídos do banco de dados. Alguns desses conhecimentos podem assumir a forma de um conjunto de regras, equações, previsões de resultados futuros, identificação de padrões, classificações etc.

Uma das tarefas cognitivas humanas mais realizadas no auxílio da compreensão do ambiente é a classificação. Carvalho (2001) afirma que essa técnica é uma das mais utilizadas nas ferramentas de *data mining*. Nós seres humanos estamos sempre classificando coisas à nossa volta. Conforme ilustra o autor, o ser humano cria classes[†] de suas relações como: colegas de trabalho, amigos, familiares, classes sociais, preconceitos etc. Para cada uma destas classes uma forma diferente de tratamento é definida. Assim, vamos criando estereótipos construídos com base no passado e prontos para serem utilizados no presente.

Da mesma maneira, os algoritmos de classificação das ferramentas de data mining buscam classificar os dados. Por exemplo, classificar os clientes em baixo, médio ou alto risco de empréstimo bancário; classificar os alunos em grande, médio ou baixo grau de evasão etc.

3 PRINCIPAIS RESULTADOS ALCANÇADOS

No decorrer da pesquisa foram aplicadas várias combinações nas variáveis do banco de dados da Universidade de Sorocaba para classificá-las e assim obter uma probabilidade estatística que classifique um determinado grupo de estudante sobre um curso.

Com a utilização desta ferramenta aplicada sobre os dados socioeconômicos/demográficos dos alunos obtivemos alguns resultados:

- Se o sexo for feminino e a faixa etária maior que 24 anos, o curso que tem maior procura é o de Pedagogia, com 26,90% das matrículas.

[†] Estrutura que abstrai um conjunto de objetos com características similares

Seminário Internacional de Educação Superior 2014

Formação e Conhecimento

Anais Eletrônicos



- Se o sexo for masculino e a faixa etária maior que 24 anos o curso que tem maior procura é o de Tecnologia em Gestão de Produção Industrial, com 13,96% de probabilidade das matrículas.

Outra análise realizada nos dados socioeconômicos/demográficos demonstrou que a cor dos estudantes não está ligada diretamente a um curso, ou seja, não há uma concentração de determinada cor em um curso.

Ao trabalhar com as questões do questionário informativo que é respondido no vestibular, foi identificado inicialmente que algumas atividades profissionais se relacionam muito com um determinado curso, como o caso da atividade industrial com o Curso Superior de Tecnologia em Gestão da Produção Industrial e professor de escola pública ou municipal com o curso de Pedagogia; ou seja, a pessoa já exerce uma profissão e vem à Universidade para se aprimorar ou progredir na carreira.

Outros resultados obtidos foram:

- Se o motivo da escolha do curso for “Valor da Mensalidade”, temos uma forte relação com a modalidade licenciatura.
- Se motivo da escolha do curso for “Ascensão profissional na atividade que exerço” há forte relação com modalidade tecnológica.

Pôde ser constatado que as pessoas que buscam ascensão profissional optam por cursos da modalidade tecnológica, por ser um curso de menor duração e muitas vezes com o valor da mensalidade mais baixo do que os de bacharelado. Já os cursos de licenciatura estão fortemente relacionados ao motivo de escolha “Valor da mensalidade”.

Após aprimorar o modelo de mineração de dados e trabalhar com outras combinações das variáveis, foi possível identificar que a pessoa que é apenas estudante e tem renda entre R\$ 1.001,00 a R\$ 2.000,00 tem uma grande probabilidade de escolher a modalidade licenciatura. Também, se a renda for maior que R\$ 2.000,00 e o motivo da escolha do curso não for “maior oferta de emprego” observa-se novamente a presença da licenciatura. Já se a renda é maior que R\$ 2.000,00 e motivo da escolha do curso for “Maior oferta de emprego”, notamos a presença da modalidade tecnológica.

Podemos entender que os estudantes que têm uma renda de até R\$ 2.000,00 optam pelos cursos da modalidade licenciatura, fato que pode estar ligado ao motivo da escolha do curso “Valor da mensalidade”.

Outra classificação identificada é que a pessoa que escolhe um curso na modalidade licenciatura e trabalha em tempo integral tende a optar por essa

Seminário Internacional de Educação Superior 2014

Formação e Conhecimento

Anais Eletrônicos



modalidade pela possibilidade de realização pessoal. Isso explica a descoberta descrita acima, na qual estudantes com renda maior que R\$ 2.000,00 não buscam o curso por maior oferta de emprego escolhendo a modalidade licenciatura.

Foi identificado também que a procura pelo curso por maior oferta de emprego está ligada à modalidade tecnológica e bacharel. Já os cursos da modalidade tecnológica são mais procurados por pessoas que buscam uma ascensão profissional em uma atividade que já desempenham.

Outra importante descoberta no cruzamento dos dados do processo seletivo com os socioeconômicos/demográficos é que as faixas etárias dos estudantes estão relacionadas com o motivo da escolha do curso. Os estudantes com até 24 anos, têm uma maior probabilidade de procurar um curso pelo motivo de realização pessoal, enquanto os demais o fazem por motivo de ascensão profissional.

Uma tendência que despertou curiosidade foi o fato dos cursos de licenciatura atraírem pessoas que o fazem objetivando realização pessoal. Já os cursos da modalidade tecnológica atraem pessoas que o buscam visando ascensão profissional ou maior oferta de emprego.

Foi compreendido, então, que os cursos tecnológicos da Universidade possuem um público que vem se qualificar profissionalmente para o mercado, enquanto nos cursos de licenciatura são pessoas que o fazem em busca da realização pessoal ou por questões financeiras.

A modalidade bacharel por ser muito ampla e possuir muitos cursos não destacou nenhuma classificação. Podemos concluir que os estudantes ali inseridos possuem motivos diversificados e distintos.

Os resultados mostram tendências que podem não se verificar em termos absolutos. Com as devidas restrições e limitações de toda pesquisa a respeito de fenômenos humanos, valem para este caso que tomou por objeto a Universidade de Sorocaba nesse período de seis anos. Embora muitas de suas conclusões possam ser observadas em outras instituições similares, não há por parte do autor nenhuma pretensão a universalizá-las. Para conclusões mais amplas e seguras seriam necessários muitos outros estudos e certamente a técnica de mineração de dados é uma ferramenta de grande utilidade.

Essa foi a primeira pesquisa empírica utilizando uma ferramenta de mineração nos dados da Universidade de Sorocaba. Uma importante contribuição para a gestão da Universidade e para o meio acadêmico, abrindo novos horizontes para pesquisas futuras.



REFERÊNCIAS

- BARATA-MOURA, J. Educação Superior: Direito ou Mercadoria?. **Avaliação**, Campinas, v. 8, n. 4, 31-36, dez. 2003.
- BERNHEIM, C. T. Desarrollo histórico de la Universidad. **Cincuentenario de la Revista Paraguaya de Sociología**, Asunción, jun. 2010. Serie Educación.
- CARVALHO, L. A. V. de. **Datamining**: a mineração de dados no marketing, medicina, economia, engenharia e administração. São Paulo: Érica, 2001. 234 p.
- CUNHA, L. A. Desenvolvimento desigual e combinado no ensino superior – estado e mercado. **Educação & Sociedade**, Campinas, v. 25, n. 88, p. 795-817, out. 2004.
- DIAS SOBRINHO, J. Educação Superior: Bem público, Equidade e Democratização. **Avaliação**: Revista da Avaliação da Educação Superior, Campinas, Sorocaba, v. 18, n. 01, 107-128, mar. 2013.
- _____. **Dilemas da educação superior no mundo globalizado**: sociedade do conhecimento ou economia do conhecimento?. São Paulo: Casa do Psicólogo, 2005. 258 p.
- _____. **Dilemas da educação superior no mundo globalizado**: sociedade do conhecimento ou economia do conhecimento?. São Paulo: Casa do Psicólogo, 2005. 258 p.
- GOERGEN, P. Da Formação ao Ensino: Um ponto cego nas políticas de Pós-Graduação. **Avaliação**: Revista da Avaliação da Educação Superior, Campinas, Sorocaba, v. 18, n. 01, 45-68, mar. 2013.
- MARQUES, W. Expansão e Oligopolização da Educação Superior no Brasil. **Avaliação**: Revista da Avaliação da Educação Superior, Campinas, Sorocaba, v. 18, n. 01, 69-84, mar. 2013.
- MELLO, H. D. A. O Banco Mundial e a Educação no Brasil: Convergências em torno de uma agenda global. 17/12/2012, p. 379. Tese Doutorado – UNICAMP.
- NAMEN, A. A.; BORGES, S.; SADALA, M. Indicadores de qualidade do ensino fundamental: o uso das tecnologias de mineração de dados e de visões multidimensionais para apoio à análise e definição de políticas públicas. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 94, n. 238, 677-700, set/dez. 2013.
- PALESTINO, C. B. **BI-business intelligence** : modelagem e tecnologia. Rio de Janeiro: Axcel Books do Brasil, 2001. 418 p.
- TURBAN, E.; RAINER, K.; POTTER, R. **Administração de tecnologia da informação**: tecnologia e prática. Rio de Janeiro: Elsevier, 2005.