



UNIVERSIDADE FEDERAL DO CEARÁ

CAMPUS DE RUSSAS

Disciplina: Introdução à Ciência da Computação.

Resumo da disciplina: Data Mining.

Discente: Paulo Henrique Diniz de Lima Alencar.

DATA MINING

A próxima equipe ficou responsável pela apresentação da cadeira de Data Mining, uma disciplina opcional ofertada pela Universidade Federal do Ceará.

O grupo iniciou sua apresentação trazendo a frase “The world is data rich but information poor” presente no livro *Data Mining: Concepts and Techniques*. A princípio a equipe buscou trazer as informações básicas sobre essa nova área, que aliás está com um gigantesco crescimento no recente âmbito social.

Após isso, a equipe buscou explicar o conceito básico do que seria Data Mining. Pelo que compreendi e tentando explicar de forma simplória, a mineração de dados seria um processo analítico que faz uso pesado da inteligência artificial, utilizado para explorar grandes quantidades de dados. Esse processo tem como intuito, localizar padrões entre as variáveis em um determinado contexto, isto é, nos grandes bancos de dados de empresas nacionais ou multinacionais, repositórios de pesquisas científicas tanto na área acadêmica ou privada, ou até mesmo na própria internet, que atualmente é uma gigantesca fonte de dados que pode gerar informações valiosíssimas.

Logo em seguinte foi explicado que a mineração de dados é dividida em várias partes, e que essas porções são divididas em outras frações, ou seja, a partir daqui podemos perceber que mineração de dados não é uma das tarefas mais simples, afinal é composta por diversas etapas que estão contidas em um enorme processo.

Como já citado anteriormente, o processo de mineração de dados é dividido em várias etapas. A divisão apresentada pela equipe que apresentou a cadeira optativas de Data Mining foi a seguinte: A seleção, que é basicamente a criação de um conjunto de dados para uso no processo da descoberta por meio da seleção dos dados. A próxima etapa é a limpeza dos dados, onde vão ser removidos os dados considerados inconsistentes. Em seguida tem-se a integração de dados, na qual várias fontes de dados

vão ser combinados de acordo com um determinado contexto. Após a integração de dados, chegamos na seleção dos dados onde os dados relevantes vão ser recuperados de um banco de dados. E por fim a transformação dos dados, onde vai ocorrer a reprojeção de alguns dados gerando outro conjunto de coordenadas que pode gerar informações de acordo os atributos recolhidos.

Dando continuidade, o grupo também explicou que os dados que são minerados podem ser classificados em duas grandes categorias, são elas: descritiva e preditiva. A descritiva caracteriza a prioridade dos dados em um grupo específicos de dados. Já a preditiva realiza uma indução dos dados dos dados atuais para fazer predições. A seguinte seria a avaliação dos padrões, logo em seguida esses padrões de dados minerados são apresentados aos usuários

Em seguida a equipe buscou responder a seguinte pergunta “O que se aprende na disciplina?” De acordo com a apresentação, em Mineração de dados analisa-se grandes quantidades de dados, buscando “separar o joio do trigo”, isto é, selecionar informações relevantes, daquilo que é considerado ruído. Depois disso, você aprende a construir relações entre esses dados que foram selecionados, com o intuito de gerar informações importantes para uma tomada de decisão. Outras técnicas que vão ser aprendidas no decorrer do curso são: associação, previsão, árvore de decisão, classificação e padrões sequencias.

Posteriormente foi respondida a seguinte pergunta “Para que o conhecimento adquirido é útil?” A equipe mostrou que no atual contexto em que vivemos, os dados crescem de maneira exponencial. Porém, “o mundo é rico em dados, mas pobre em informações”, então a partir dessa frase, fica notório que atualmente é imprescindível o uso desse tipo de tecnologia em determinados campos, buscando resolver o problema da dificuldade da análise de milhares de dados em conjunto.

E para terminar, com relação aos professores que podem dar a disciplina de Data Mining, o grupo obteve as seguintes informações em relação aos campos de Russas, Quixadá e Pici. Em relação ao Campus de Russas, a professora Dr. Marilha Soares Nunes Mari possui capacidade de ministra essa cadeira. Em relação aos outros campuses da UFC, a equipe não disponibilizou nenhuma informação sobre isso durante a apresentação.