



UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE
COMPUTAÇÃO

Departamento de Ciências de Computação

Disciplina: SCC0630 – Inteligência Artificial

Professora: Solange Oliveira Rezende

Estagiária PAE: Vitor Rodrigues Tonon

Trabalho Prático

1. Objetivo do Trabalho:

Com o crescimento e a maior visibilidade da profissão “Cientista de Dados”, esse trabalho visa proporcionar a experiência de mineração de dados desde o levantamento e preparação dos dados até a exploração e análise dos resultados.

2. Especificação:

O trabalho é dividido em 3 partes: 1) Levantamento e coleta de dados. 2) Proposta de trabalho e detalhamento da mineração. 3) Aplicação da mineração de dados e análise dos resultados obtidos.

1) O Levantamento de dados consiste em encontrar e coletar uma base de dados da escolha do aluno. Essa base de dados deve ser enviada ao Tidia junto com um relatório, que descreve a base, seus campos e os possíveis conhecimentos úteis que podem ser obtidos através dela. Caso a base seja muito grande, entrar em contato com o estagiário PAE para definir outra forma de entrega.

2) A proposta consiste em um documento detalhado, propondo eventuais pré-processamentos a base, técnicas a serem utilizadas e objetivo de exploração. É de extrema importância que as técnicas escolhidas sejam justificadas de acordo com o objetivo de exploração.

3) A última etapa consiste na mineração da base de dados utilizando as técnicas propostas e análise de conhecimento obtido. Nessa última etapa o grupo deverá fazer um relatório, descrevendo a mineração de dados, eventuais mudanças que foram necessárias, problemas e soluções encontradas e a experiência que foi adquirida com o trabalho. O relatório também deve ter uma discussão dos resultados obtidos, se foram satisfatórios (e esperados) ou não. No final, um pôster deverá ser feito explicando as 3 etapas de maneira sucinta, descrevendo também os resultados obtidos.

3. Avaliação:

O trabalho será avaliado com base na média simples dos relatórios (RM) e na apresentação (A) do trabalho.

Para RM e A maiores que 5,0 a Nota do Trabalho (NT) será calculada da seguinte forma:

$$NT = RM \cdot 0,6 + A \cdot 0,4$$

Caso contrário:

$NT = \text{MenorNota}(RM, A)$

4. **Entrega:**

O trabalho deverá ser entregue **até o dia 10/06** na atividade correspondente no Tidia. Criar um zip no seguinte formato: T1_GrupoX..zip, contendo: (i) base de dados utilizada, (ii) códigos utilizados na mineração, (iii) relatório **em PDF**, (iv) apresentação **em PDF**.

5. **Informações Adicionais:**

5.a) Relatório

O objetivo do relatório é apresentar o que foi feito e deve conter todas as informações relativas ao trabalho.

Parte 1) Deverá descrever a base que será utilizada, seus campos e as possibilidades de mineração de dados (qual conhecimento é interessante extrair).

Parte 2) Deverá descrever todo o processamento que será realizado na base de dados, desde o pré-processamento, até qual(is) técnica(s) será(ão) utilizada(s) com o objetivo da mineração. É extremamente importante que as escolhas sejam justificadas e o objetivo de exploração esteja bem definido.

Parte 3) Deverá descrever o processo de mineração de dados, eventuais mudanças que foram necessárias no planejamento, problemas e soluções encontradas e a experiência que foi adquirida com o trabalho. O relatório também deve ter uma discussão dos resultados obtidos, se foram satisfatórios (e esperados) ou não.

5.b) Apresentação

- Criar slides para apresentar o trabalho desenvolvido, contendo: descrição do problema e da base de dados utilizada; pré-processamento realizado; algoritmos de AM utilizados e resultados obtidos. A apresentação será no formato de um workshop, a ser realizado no dia 11/06.
- A escolha do integrante do grupo a apresentar o trabalho será aleatória. Portanto, estejam todos preparados.
- Na apresentação serão avaliadas a organização das informações no pôster, a apresentação do trabalho realizado e a postura dos integrantes do grupo.