



Universidade Federal de Alagoas  
Instituto de Computação  
Ciência de Dados



## Projeto

### Proposta

Aplicação de Ciência de Dados em informações sobre gestão de pessoas do Poder Executivo Federal disponibilizadas pelo Ministério da Economia.

Deve-se analisar conjuntamente:

- Informações sobre abono de permanência. Descrição e bases de dados no link: <http://www.dados.gov.br/dataset/gastos-pessoal-abono-permanencia>
- Informações sobre aposentados. Descrição e bases de dados no link: <http://www.dados.gov.br/dataset/aposentados-poder-executivo-federal>

### Conteúdo

O arquivo de texto do projeto deve apresentar, pelo menos, as seguintes seções:

- **Aplicação:** detalhamento da proposta do trabalho;
- **Experimentos**
  - **Base de dados:** descrição da origem dos dados, técnicas de pré-processamento (limpeza, integração, redução);
  - **Estatística descritiva e inferência:** uso de gráficos e estatísticas para detalhar os dados;
  - **Métodos avaliados:** descrição de métodos de Aprendizagem de Máquina usados;
  - **Métricas de avaliação:** métricas usadas para avaliar os métodos (F-measure, AUC, taxa de acerto);
  - **Métodos de avaliação:** uso de amostragem para avaliar os modelos (k-fold cross-validation, leave-one-out, bootstrap)
- **Resultados:** apresentação dos resultados e discussão sobre as observações. Fazer uso de testes de hipótese para dar suporte às afirmações.

- **Conclusão:** resumo do que foi realizado no trabalho e principais observações.
- **Bibliografia:** obras consultadas ou citadas.

## **Estilo**

O documento deve seguir o modelo de confecção de artigos para conferências segundo a Sociedade Brasileira de Computação (SBC). O link [1] possui modelos em doc e latex do site da SBC. O link [2] possui o modelo em latex da Overleaf.

[1]<https://www.sbc.org.br/documentos-da-sbc/summary/169-templates-para-artigos-e-capitulos-de-livros/878-modelosparapublicaodeartigos>

[2]<https://pt.overleaf.com/latex/templates/sbc-conferences-template/blbxwjwzdngr>

O tamanho é limitado a 12 páginas.

## **Apresentação**

O aluno deve apresentar o seu trabalho motivando sua análise, metodologia de tratamento dos dados, metodologia de experimentação e conclusões.

## **Submissão**

Deve-se entregar os seguintes arquivos (em .zip ou .rar):

- Documento do texto (em pdf);
- Código utilizado nos experimentos;
- Apresentação;

Enviar usando o classroom da disciplina na atividade “Projeto”.

## **Avaliação**

- Uso abrangente e adequado de Estatística, Mineração de Dados, Aprendizagem de Máquina, Visualização de Dados (4,0 pontos);
- Uso das informações obtidas para aplicação à tomada de decisão (6,0 pontos).