

Trabalho Final – Aplicação de Técnicas de Aprendizado de Máquina em Dados Públicos Brasileiros

- O objetivo deste trabalho final é aplicar, de forma prática, os conhecimentos adquiridos ao longo da disciplina de Inteligência Artificial (IA) na resolução de problemas reais a partir de bases de dados públicas brasileiras.
- Os grupos deverão escolher um **conjunto de dados de domínio público**, dentro de áreas de relevância social (como saúde, educação, energia, meio ambiente, segurança pública ou áreas afins).
- Os grupos deverão desenvolver modelos de **regressão ou classificação**, conforme a temática sorteada em sala de aula.

Trabalho Final - Orientações gerais

- Os trabalhos serão desenvolvidos por equipes de até 03 pessoas
- Cada grupo deverá implementar no mínimo três algoritmos de aprendizado de máquina abordados na disciplina, realizar um benchmarking entre os modelos utilizando métricas de avaliação de desempenho adequadas ao tipo de tarefa (classificação ou regressão) e, com base nos resultados, avaliar qual modelo apresenta melhor desempenho para o problema proposto
- O trabalho final deve ser entregue em formato de artigo científico
- Além da implementação técnica dos algoritmos, espera-se que o trabalho reflita criticamente sobre os dados utilizados e os possíveis impactos sociais da solução desenvolvida
- **Não serão aceitos trabalhos que NÃO utilizem base de dados públicos brasileiros**

Trabalho Final – Áreas Temáticas

- Os grupos deverão escolher um conjunto de dados de domínio público, dentro de áreas de relevância social, tais como:
- Saúde
- Educação
- Energia
- Meio Ambiente
- Segurança Pública
- Ou áreas afins (desde que utilizem dados públicos brasileiros)

Trabalho Final – Áreas Temáticas

Sugestões de bases públicas brasileiras na área da Saúde



- **DATASUS / TABNET (Ministério da Saúde)**
- Link: <https://datasus.saude.gov.br>
- Dados sobre internações, mortalidade, doenças de notificação compulsória, vacinação, etc.
- Exemplo de potenciais estudos: predição de surtos, análise de mortalidade, acesso a serviços, etc.

- **OpenDataSUS (Ministério da Saúde)**
- Link: <https://opendatasus.saude.gov.br>
- Dados de Vigilância em Saúde, notificação de doenças, imunização, mortalidade, hospitalização e internações, estabelecimentos de saúde, atenção básica, etc.
- Exemplo de potenciais estudos: Predição de casos de doenças com base em dados regionais, análise do risco de mortalidade em pacientes com comorbidades específicas, regressão para estimar tempo médio de internação ou custos por tipo de procedimento, etc.

Trabalho Final – Áreas Temáticas

Sugestões de bases públicas brasileiras na área de Energia

- **Agência Nacional de Energia Elétrica (ANEEL)**
 - Link: <https://dadosabertos.aneel.gov.br>
 - Dados sobre geração distribuída, consumo energético, tarifas, outorgas.
 - Possíveis estudos: predição de consumo, análise de desigualdade no acesso à energia, etc.
-
- **ONS (Operador Nacional do Sistema Elétrico)**
 - Link: <https://dados.ons.org.br>
 - Dados sobre carga, geração e intercâmbio de energia.
 - Possíveis estudos: Previsão da geração de energia por fonte, Previsão de carga energética, Classificação de dias de alto risco de sobrecarga, etc.



Trabalho Final – Áreas Temáticas

Sugestões de bases públicas brasileiras área de **Meio Ambiente**

- **IBAMA – Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis**
- Link: <https://dadosabertos.ibama.gov.br/dataset/>
- Dados sobre SINAFLOR, arrecadações, etc.
- **IBAMA: Cadastro Técnico Federal (CTF)**
- Link: <https://dadosabertos.ibama.gov.br/dataset/pessoas-juridicas-inscritas-no-ctf-app>
- Relação das pessoas jurídicas que efetuaram a inscrição no Cadastro Técnico Federal de Atividades Potencialmente Poluidoras e Utilizadoras de Recursos Naturais – CTF/APP.
- Exemplos de potenciais estudos: classificar tipos de atividades econômicas mais associadas a infrações ambientais, classificar regiões mais vulneráveis a desmatamento ilegal, Prever o valor médio de multas ambientais com base em tipo de infração, localização, e reincidência, etc.
- **INPE – Instituto Nacional de Pesquisas Espaciais**
- **Dados de Desmatamento (sistemas PRODES, DETER)**
- Link: <https://terrabrasilis.dpi.inpe.br>
- Dados de desmatamento na Amazônia e no Cerrado (análises temporais, geoespaciais, etc).



Trabalho Final – Áreas Temáticas

Sugestões de bases públicas brasileiras na área da Educação

- **INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira)**
- **Censo Escolar / ENEM / Prova Brasil / SAEB**
- Link: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos>
- Dados do Censo Escolar da Educação Básica, Censo da Educação Superior, Resultados de Avaliações Nacionais, Indicadores educacionais, Microdados detalhados.
- Possíveis estudos: predição de notas do estudante a partir de variáveis socioeconômicas, previsão de indicadores educacionais, previsão de evasão escolar, classificação de risco de evasão escolar, etc
- **QEdu (dados educacionais estruturados)**
- Link: <https://qedu.org.br>
- Interface mais amigável para os dados do INEP.



Trabalho Final – Áreas Temáticas

Sugestões de bases públicas brasileiras na **Área de Segurança Pública**

- **SINESP - mações de Segurança Pública, Prisionais e sobre Drogas**
- **Ministério da Justiça**



- Dados nacionais de segurança pública, crimes, UFs, Municípios, etc.
- Link: <https://dados.gov.br/dados/conjuntos-dados/sistema-nacional-de-estatisticas-de-seguranca-publica>
- **Fórum Brasileiro de Segurança Pública (FBSP)**
- Link: <https://forumseguranca.org.br>
- Relatórios anuais e datasets sobre homicídios, feminicídios, violência policial.
- Potenciais estudos: Classificação de municípios com alto risco criminal, previsão de feminicídio com base em fatores de risco, previsão do número de homicídios por município, previsão do crescimento da população carcerária, etc.

Trabalho Final – Áreas Temáticas

Portal Brasileiro de Dados Abertos e Catálogo Nacional de Dados!

Link: <https://dados.gov.br/home>

Nele é possível encontrar dados publicados pelo governo federal e por governos locais para realizar pesquisas, desenvolver aplicativos e criar novos serviços.



14.481

Conjuntos de dados abertos



93.183

Recursos

Trabalho Final - Orientações sobre o artigo

- O artigo científico deverá seguir o modelo da Sociedade Brasileira de Computação (Modelos para publicação de artigos disponível em <<https://www.sbc.org.br/documentosinstitucionais/#publicacoes>>).
- Os discentes terão de 20 a 30 minutos para apresentar seu trabalho para a turma nos dias **08 e 11 de julho***, no horário da aula, conforme sorteio em sala de aula (trabalhos de classificação no dia 08/07, trabalhos de regressão no dia 11/07).
- *Há uma possibilidade de explorar uma data no mês de agosto. (A confirmar)
- **Além do artigo, os alunos deverão anexar o código utilizado para o treinamento e avaliação dos modelos.**
- A nota final dos discentes será composta a partir da avaliação do trabalho escrito (artigo) e apresentação oral (presencial).

Trabalho Final - Instruções para o artigo (de 08 a 10 páginas – sem contar anexo)

- O artigo deverá ser obrigatoriamente composto dos seguintes tópicos:
 1. **Título:** Deve refletir claramente o objetivo ou a principal aplicação do trabalho.
 2. **Resumo:** Deve incluir uma breve descrição do problema abordado, os algoritmos utilizados, os principais resultados e as conclusões do estudo. (em PT e EN)
 3. **Palavras-chave:** Selecione cinco palavras-chave que reflitam os principais conceitos abordados no artigo.
 4. **Introdução:** Esta seção deve contextualizar o problema, explicando sua importância e relevância na área. Além disso, deve apresentar claramente o objetivo da pesquisa ou aplicação. Finalize com uma visão geral da estrutura do artigo.

Trabalho Final - Instruções para o artigo (de 08 a 10 páginas – sem contar anexo)

5. Trabalhos Relacionados: Apresente pesquisas e soluções anteriores relacionadas ao problema abordado, incluindo técnicas de aprendizado de máquina já aplicadas no contexto do trabalho e os resultados obtidos. Analise como o seu trabalho se relaciona com os trabalhos desta seção.

6. Fundamentação Teórica (ou Referencial Teórico): Apresente de maneira clara e sucinta os conceitos base que embasam seu trabalho (por exemplo: descreva os algoritmos de IA utilizados no trabalho e seus principais parâmetros)

Trabalho Final - Instruções para o artigo (de 08 a 10 páginas – sem contar anexo)

7. Metodologia: Descreva detalhadamente como o problema ou tarefa foi resolvido. Inclua informações sobre o dataset escolhido, suas características e fontes. Descreva também as etapas de preparação dos dados, como limpeza, normalização ou balanceamento, se necessário. Justifique a escolha dos algoritmos de aprendizado de máquina utilizados com base nas características do problema. Apresente a configuração experimental, incluindo os parâmetros dos modelos, as métricas de avaliação e as técnicas de validação (validação cruzada, etc), se aplicável.

Trabalho Final - Instruções para o artigo (de 08 a 10 páginas – sem contar anexo)

8. Resultados numéricos e discussão:

Compare os resultados obtidos pelos diferentes algoritmos utilizando as métricas escolhidas (por exemplo: acurácia, precisão, recall, erro quadrático médio, conforme a natureza do problema). Utilize visualizações, como tabelas de comparação, gráficos de desempenho e matrizes de confusão, para facilitar a compreensão dos resultados. Fale das vantagens e limitações dos algoritmos escolhidos.

Interprete os resultados obtidos dentro do contexto do trabalho. Não basta apresentar valores de métricas ou gráficos. É necessário analisar o que esses resultados realmente significam em relação ao problema estudado. Compare com trabalhos relacionados (se possível).

Trabalho Final - Instruções para o artigo (de 08 a 10 páginas – sem contar anexo)

9. Considerações Finais

Resuma os principais achados do trabalho, destacando como os resultados contribuem para a área de estudo. Identifique as limitações do trabalho realizado e sugira possíveis direções para futuras pesquisas ou melhorias.

10. Referências Bibliográficas

Liste todas as fontes citadas no trabalho, de acordo com o formato de citação recomendado pelas normas da ABNT.

Apêndices ou Anexos

Inclua o código do estudo ou qualquer outro material relevante, como dados complementares ou detalhes da implementação.

Trabalho Final - Instruções para a apresentação do trabalho

- Os slides devem refletir os principais pontos do trabalho (siga a mesma organização lógica do artigo).
- Não deve conter mais do que 20 slides.
- **Menos é mais!**
 - Use fundo claro com texto escuro (ou vice-versa).
 - Escolha fontes legíveis (ex: Arial, Calibri).
 - Evite poluição visual (muitos elementos, efeitos ou imagens desnecessárias).
- O tempo de apresentação é de 20 a 30 minutos. Uma parte desse tempo (ao final da apresentação) será usada para mostrar os principais pontos do código.

Trabalho Final - Entregáveis

- Ao todo, há 03 entregáveis a serem encaminhados via SIGAA, conforme prazo cadastrado:
 - Artigo científico
 - Apresentação de slides
 - Código

Cr terios para a avalia  o da apresenta  o oral do grupo

Os seguintes cr terios s o considerados:

- Contextualiza  o do trabalho, apresenta  o da problem tica (20%)
- Compreens o dos conceitos e da aplica  o dos algoritmos no trabalho (30%);
- Clareza na explica  o de cada slide (20%);
- Qualidade visual e organiza  o dos slides (10%);
- Cumprimento do tempo estabelecido (10%);
- Participa  o equilibrada dos membros da equipe (10%).

Grupos

1. Flávia, Yorrán e Alexandre (Classificação)
2. Vinicius, Emilly, Isabelly (Classificação)
3. Daniel, Jamir, César (Classificação)
4. Evandro, Iago e Antonio Lucas (Regressão)
5. Reynaldo, Stela e Lenon (Regressão)
6. Hugo, Klíssia, Gustavo (Regressão)
7. Klévia, Marcos, Diego (Classificação)
8. Maria, João, Lucas (Classificação)
9. Elias (Classificação)
10. Rodrigo (Regressão)
11. André Gabriel (Regressão)

Atenção!

- Os arquivos referentes ao trabalho (artigo, código e slides) deverão ser entregues via Sigaa até **07/07***. **NÃO SERÃO ACEITOS TRABALHOS APÓS ESTE PRAZO.**
- ***Irei confirmar se haverá possibilidade de transferir para agosto.**
- **Trabalhos não encaminhados no prazo não poderão ser apresentados.**
- Os trabalhos de classificação serão apresentados no dia 08/07 (terça-feira) e os trabalhos de regressão serão apresentados no dia 11/07 (sexta-feira)



Dúvidas??

Obrigada 😊