



FACULDADE CATÓLICA DA PARAÍBA
Instituição Mantida pela Fundação de Ensino Superior de Cajazeiras

NÚCLEO DE EXTENSÃO E PESQUISA ACADÊMICA – NEPA
PROPOSTA DE PROJETO DE PESQUISA
EDITAL N.º 01/2024

Documentos indispensáveis para inscrição

- 1 (uma) cópia por e-mail do Formulário de **Proposta de Projeto de Extensão**
- 1 (uma) cópia por e-mail do Currículo Lattes atualizado.

1. Nome do(a) professor(a) proponente do Projeto: Yuri Rennan Campos Silva
2. Titulação do Professor Proponente: <input checked="" type="checkbox"/> Especialista <input type="checkbox"/> Mestre <input type="checkbox"/> Doutor <input type="checkbox"/> Pós Doutor
2.1. Telefones para contato: (83) 981399979
2.2 E-mail: yurirennan14@gmail.com

3. Curso: Ciências da Computação

4. Título do Projeto: Mineração de Dados: Avaliação de Atributos socioeconômicos da municípios da Mesorregião de Cajazeiras

5. Linha de Pesquisa: Inteligência Computacional
--

6. Situação do Projeto: <input checked="" type="checkbox"/> Novo <input type="checkbox"/> em andamento
--

7. Resumo: <p>O presente projeto de pesquisa propõe a aplicação de técnicas de mineração de dados para avaliar os atributos socioeconômicos dos municípios localizados na Mesorregião de Cajazeiras. A região em foco apresenta uma diversidade significativa em termos de características socioeconômicas, abrangendo desde áreas urbanas densamente povoadas até regiões rurais com menor desenvolvimento econômico. A mineração de dados emerge como uma ferramenta poderosa para analisar grandes conjuntos de dados e extrair informações relevantes para a compreensão dessas disparidades socioeconômicas.</p> <p>Este estudo visa identificar padrões, tendências e correlações nos dados socioeconômicos dos municípios da Mesorregião de Cajazeiras, utilizando técnicas de mineração de dados como clustering, classificação e análise de associação. A análise dos dados permitirá uma compreensão mais profunda das características socioeconômicas dos municípios, destacando fatores que influenciam no desenvolvimento local, como renda per capita,</p>



FACULDADE CATÓLICA DA PARAÍBA
Instituição Mantida pela Fundação de Ensino Superior de Cajazeiras

índices de educação, saúde, emprego e infraestrutura.

Os resultados esperados deste projeto incluem a identificação de padrões de desenvolvimento socioeconômico, a detecção de áreas com necessidades específicas de intervenção e o fornecimento de insights para políticas públicas voltadas para o desenvolvimento regional. Além disso, este estudo contribuirá para o avanço da aplicação de técnicas de mineração de dados em contextos socioeconômicos, demonstrando seu potencial na análise e compreensão de complexos conjuntos de dados regionais.

8. Palavras-chave (Indique até 3 palavras que melhor caracterizem a proposta)

Indicadores Socioeconômicos

Mineração de dados

Mesorregião de Cajazeiras

9. Justificativa

A Mesorregião de Cajazeiras, situada no estado da Paraíba, apresenta uma diversidade socioeconômica marcante, refletida em disparidades significativas entre seus municípios. Essa heterogeneidade demanda uma compreensão detalhada das condições locais para a formulação e implementação eficaz de políticas públicas e estratégias de desenvolvimento regional. Nesse contexto, a aplicação de técnicas de mineração de dados surge como uma abordagem promissora para analisar e extrair insights valiosos a partir dos vastos conjuntos de dados disponíveis sobre a região.

A justificativa para este projeto se fundamenta na necessidade de uma análise abrangente e sistemática dos atributos socioeconômicos dos municípios da Mesorregião de Cajazeiras. Através da mineração de dados, será possível identificar padrões ocultos, correlações e tendências que podem não ser facilmente perceptíveis por meio de métodos tradicionais de análise. Isso permitirá uma compreensão mais holística das dinâmicas socioeconômicas locais, contribuindo para a tomada de decisões mais informadas por parte dos gestores públicos, acadêmicos e demais interessados na região.

Além disso, a aplicação de técnicas de mineração de dados neste contexto oferece uma oportunidade única de explorar a riqueza dos dados disponíveis e transformá-los em conhecimento acionável. Ao identificar áreas de maior vulnerabilidade, padrões de desenvolvimento e necessidades específicas de cada município, este projeto poderá fornecer subsídios importantes para a formulação de políticas públicas mais eficazes, direcionadas às demandas reais da população local.

Portanto, diante da complexidade e relevância dos desafios socioeconômicos enfrentados pela Mesorregião de Cajazeiras, este projeto se justifica como uma iniciativa crucial para promover o desenvolvimento regional sustentável, baseado em evidências e orientado para



FACULDADE CATÓLICA DA PARAÍBA
Instituição Mantida pela Fundação de Ensino Superior de Cajazeiras

resultados tangíveis.

10. Fundamentação Teórica

A fundamentação teórica deste projeto se baseia em conceitos e abordagens da mineração de dados, bem como em teorias socioeconômicas relevantes para a compreensão dos fenômenos observados nos municípios da Mesorregião de Cajazeiras.

1. **Mineração de Dados (Data Mining):** A mineração de dados é uma disciplina interdisciplinar que combina técnicas e métodos de áreas como estatística, inteligência artificial, aprendizado de máquina e banco de dados. Seu objetivo é descobrir padrões, tendências e relações significativas em grandes conjuntos de dados. Neste projeto, serão aplicadas técnicas de mineração de dados, como clustering, classificação e análise de associação, para explorar e extrair informações relevantes dos dados socioeconômicos dos municípios da Mesorregião de Cajazeiras.
2. **Teoria Socioeconômica:** Esta abordagem compreende teorias e conceitos que explicam as interações entre fatores socioeconômicos em uma determinada região. Teorias como o desenvolvimento econômico local, capital social, teorias do crescimento endógeno, e teorias sobre o papel das instituições na promoção do desenvolvimento regional podem fornecer insights valiosos para entender as dinâmicas socioeconômicas observadas na Mesorregião de Cajazeiras. Através dessa base teórica, será possível interpretar os resultados da mineração de dados à luz de conceitos socioeconômicos fundamentais.
3. **Indicadores Socioeconômicos:** Os indicadores socioeconômicos são medidas quantitativas que refletem aspectos da qualidade de vida, bem-estar e desenvolvimento de uma determinada população. Estes indicadores incluem, entre outros, dados sobre renda per capita, educação, saúde, emprego, infraestrutura e desigualdade econômica. Ao selecionar e analisar cuidadosamente esses indicadores, é possível obter uma visão abrangente das condições socioeconômicas dos municípios da Mesorregião de Cajazeiras.
4. **Geografia Econômica e Desenvolvimento Regional:** Estudos na área de geografia econômica e desenvolvimento regional oferecem insights sobre



FACULDADE CATÓLICA DA PARAÍBA
Instituição Mantida pela Fundação de Ensino Superior de Cajazeiras

os fatores que influenciam o desenvolvimento socioeconômico em nível local e regional. Teorias sobre aglomerações industriais, economias de escala, efeitos de vizinhança e migração populacional podem fornecer um quadro conceitual para entender as disparidades socioeconômicas observadas na Mesorregião de Cajazeiras e orientar a interpretação dos resultados da mineração de dados.

Por meio da integração dessas perspectivas teóricas, este projeto busca não apenas analisar os dados socioeconômicos dos municípios da Mesorregião de Cajazeiras, mas também interpretar esses dados à luz de teorias socioeconômicas fundamentais. Essa abordagem teórica integrada permitirá uma compreensão mais profunda das dinâmicas socioeconômicas locais e fornecerá insights valiosos para a formulação de políticas públicas e estratégias de desenvolvimento regional.

11. Objetivo Geral

O objetivo geral deste projeto de pesquisa é aplicar técnicas de mineração de dados para avaliar os atributos socioeconômicos dos municípios da Mesorregião de Cajazeiras, visando identificar padrões, tendências e correlações que possam fornecer insights para o desenvolvimento regional e a formulação de políticas públicas.

12.: Objetivos Específicos:

1. Coletar dados socioeconômicos abrangentes dos municípios da Mesorregião de Cajazeiras, incluindo indicadores relacionados à renda, educação, saúde, emprego, infraestrutura e desigualdade econômica.
2. Preparar e pré-processar os dados para análise, realizando limpeza, integração e transformação dos conjuntos de dados brutos, garantindo a qualidade e a consistência dos dados utilizados.
3. Aplicar técnicas de mineração de dados, como clustering, classificação e análise de associação, para identificar padrões e agrupamentos nos dados socioeconômicos dos municípios, permitindo uma compreensão mais profunda das dinâmicas regionais.
4. Analisar e interpretar os resultados da mineração de dados, identificando correlações entre os diferentes atributos socioeconômicos e destacando áreas de maior vulnerabilidade ou potencial de desenvolvimento em cada município.



FACULDADE CATÓLICA DA PARAÍBA
Instituição Mantida pela Fundação de Ensino Superior de Cajazeiras

5. Elaborar relatórios e visualizações dos resultados obtidos, fornecendo subsídios para a formulação de políticas públicas e estratégias de desenvolvimento regional baseadas em evidências.
6. Contribuir para o avanço do conhecimento na aplicação de técnicas de mineração de dados em contextos socioeconômicos, demonstrando o potencial dessas abordagens para a análise e compreensão de complexos conjuntos de dados regionais.

13. Metodologia:

1. Coleta de Dados:

- Realizar levantamento de dados socioeconômicos dos municípios da Mesorregião de Cajazeiras a partir de fontes confiáveis, como o Instituto Brasileiro de Geografia e Estatística (IBGE), o Instituto de Pesquisa Econômica Aplicada (IPEA) e órgãos governamentais estaduais e municipais.
- Coletar informações sobre indicadores socioeconômicos relevantes, incluindo renda per capita, nível de educação, acesso à saúde, taxa de emprego, infraestrutura urbana e desigualdade econômica.

2. Pré-processamento de Dados:

- Realizar a limpeza dos dados, identificando e tratando valores ausentes, inconsistências e outliers.
- Integrar diferentes conjuntos de dados e formatos para garantir a consistência e a qualidade dos dados utilizados na análise.
- Realizar transformações necessárias nos dados, como normalização e codificação de variáveis categóricas, para prepará-los para as etapas seguintes da análise.

3. Análise Exploratória de Dados:

- Realizar análises estatísticas descritivas para explorar a distribuição e a variabilidade dos dados socioeconômicos em cada município.
- Identificar padrões e tendências preliminares nos dados através de visualizações gráficas, como histogramas, gráficos de dispersão e mapas temáticos.

4. Aplicação de Técnicas de Mineração de Dados:



FACULDADE CATÓLICA DA PARAÍBA
Instituição Mantida pela Fundação de Ensino Superior de Cajazeiras

- Utilizar técnicas de mineração de dados, como clustering, classificação e análise de associação, para explorar os padrões e agrupamentos nos dados socioeconômicos dos municípios.
- Aplicar algoritmos de clustering para identificar grupos de municípios com características socioeconômicas semelhantes.
- Realizar análises de classificação para prever ou categorizar variáveis socioeconômicas com base em outras variáveis.
- Investigar relações entre diferentes variáveis socioeconômicas através de técnicas de análise de associação.

5. Interpretação dos Resultados:

- Analisar e interpretar os resultados da mineração de dados, identificando padrões significativos e correlações entre os atributos socioeconômicos dos municípios.
- Identificar áreas de maior vulnerabilidade ou potencial de desenvolvimento com base nos padrões identificados.
- Avaliar a relevância dos resultados para a formulação de políticas públicas e estratégias de desenvolvimento regional.

6. Elaboração de Relatórios e Visualizações:

- Elaborar relatórios técnicos detalhando os métodos utilizados, os resultados obtidos e as conclusões da análise.
- Criar visualizações gráficas claras e informativas, como gráficos de barras, mapas e diagramas de dispersão, para comunicar os resultados de forma acessível a diferentes públicos.

7. Validação e Discussão dos Resultados:

- Validar os resultados obtidos através de análises adicionais e comparações com estudos anteriores ou dados históricos.
- Discutir as limitações do estudo e possíveis fontes de viés nos resultados.
- Explorar implicações práticas e teóricas dos resultados para o desenvolvimento regional e a formulação de políticas públicas.

Essa metodologia integrada permitirá uma análise abrangente dos atributos socioeconômicos dos municípios da Mesorregião de Cajazeiras, fornecendo insights valiosos para o desenvolvimento regional e a tomada de decisões informadas pelos gestores públicos e demais interessados na região.



FACULDADE CATÓLICA DA PARAÍBA
Instituição Mantida pela Fundação de Ensino Superior de Cajazeiras

14. Cronograma de atividades (Informar as atividades a serem executadas)

1. Preparação e Coleta de Dados
2. Pré-Processamento de Dados
3. Análise Exploratória de Dados
4. Aplicação de Técnicas de Mineração de Dados
5. Interpretação dos Resultados
6. Elaboração dos Resultados e Visualizações
7. Validação e Discussão dos Resultados
8. Escrita de Artigo científico para publicação

Mês atividades	FEV	MAR	ABR	MAI	JUN	AGO	SET	OUT	NOV	DEZ
	1	1	2	3	3	4	4	6	6	7
	2	2	3	4	4	5	5	7	7	8
	8	8	8	8	5	6	6	8	8	
					8	8	8			

15. Referências:

AMO, S. Técnicas de mineração de dados. Universidade Federal de Uberlândia, Faculdade de Computação. Disponível em https://www.researchgate.net/profile/Sandra_Amo/publication/260300816_Tecnicas_de_Minerao_de_Dados/links/54230bd80cf290c9e3_ae25e3/Tecnicas-de-Minerao-de-Dados.pdf , Acesso em: 10 de Maio de 2022.

AZANK, Felipe. Como avaliar seu modelo de regressão : as principais métricas para avaliar seus modelos de regressão. Medium. Disponível em: <https://medium.com/turing-talks/como-avaliar-seu-modelo-de-regress%C3%A3o-c2c8d73dab96> . Acesso em: 16 de maio de 2022

BROWNLLE, Jason. Como usar as transformações StandardScaler e MinMaxScaler em Python. Domínio do aprendizado de máquina, 2020. Disponível em: <https://machinelearningmastery.com/standardscaler-and-minmaxscaler-transforms-in-python/> .Acesso em: 19 de e maio de 2022.

FAYYAD, U; PIATETSKY-SHAPIRO, G; SMYTH, P. From Data Mining to Knowledge Discovery and Data Mining: Towards a Unifying Framework, 1996.

FIGUEIREDO, Dalson Brito; et al. Desvendando os Mistérios do Coeficiente de Correlação de Pearson: O retorno. Leviathan | Caderno de Pesquisa Política, São Paulo,



FACULDADE CATÓLICA DA PARAÍBA
Instituição Mantida pela Fundação de Ensino Superior de Cajazeiras

N. 8, pp. 66- 95, 2014.

FONSECA, J. J. S. Metodologia da pesquisa científica. Fortaleza: UEC, 2002. Apostila.

GUIMARÃES, Alysson. Seleção de Recursos com Eliminação de Recursos Recursivos (RFE). Medium, 2021. Disponível em:

<https://k3ybladewielder.medium.com/sele%C3%A7%C3%A3o-de-features-com-recursive-feature-elimination-rfe-5effad69590b>. Acesso em: 19 de e maio de 2022.

HOPPEN, Joni; PRATES; Wladimir. Outliers, o que são e como tratá-los em uma análise de dados? Aquarela, 2022. Disponível em: <https://www.aquare.la/o-que-sao-outliers-e-como-trata-los-em-uma-analise-de-dados/>. Acesso em: 19 de e maio de 2022.

Lima, Milson Louseiro. Um Modelo para Predição de Bolsa de Valores Baseado em Mineração de Opinião, São Luís -MA: 2016. 113 p.

Painel de indicadores. Instituto Brasileiro de Geografia e Estatística-IBGE. Disponível em: <https://www.ibge.gov.br/indicadores>. Acesso em: 15 de maio de 2022.

RODRIGUES, Rodrigo Lins. Uma abordagem de Mineração de Dados Educacionais para previsão de desempenho a partir de padrões comportamentais de Autorregulação da Aprendizagem. Tese de Doutorado. Centro de Informática/UFPE. Programa de Pós Graduação em Ciência da Computação da Universidade Federal de Pernambuco (UFPE). 2016.

RUGARD, Dieter Siedenberg. Indicadores de desenvolvimento socioeconômico uma síntese. Desenvolvimento em questão, Rio Grande do Sul, vol. 1, núm. 1, pp. 45 - 71, Janeiro/Junho, 2003. Todas as pesquisas e estudos. Instituto Brasileiro de Geografia e Estatística-IBGE. Disponível em: <https://www.ibge.gov.br/estatisticas/todos-os-produtos-estatisticas.html>. Acesso em: 15 de maio de 2022.

ZEVIANI, Walmes Marques; JR, Paulo Justiniano Ribeiro; BONAT, Wagner Hugo. Modelos de Regressão Não Linear. UFP. Disponível em: http://www.leg.ufpr.br/~walmes/cursoR/mrnl_bookdown/modelos-nao-lineares.html . Acesso em: 17 de e maio de 2022.

ZEVIANI, Walmes Marques; JÚNIOR, Paulo Justiniano Ribeiro; BONAT, Wagner Hugo. Modelos de regressão não linear. Paraná: 58° RBRAS e 15° SEAGRO, 2013.



FACULDADE CATÓLICA DA PARAÍBA
Instituição Mantida pela Fundação de Ensino Superior de Cajazeiras

16. Termo de compromisso do solicitante (preenchimento obrigatório)

Declaro, para fins de direito, conhecer as normas gerais fixadas no Edital N° 01/2024 para chamada de Projetos de Pesquisa e Extensão – 2024, da Faculdade Católica da Paraíba e assumo o compromisso de dedicar-me às atividades durante a vigência do projeto.

Local

Cajazeiras, PB

Data

19/02/2024

Assinatura

Luiz Rômulo C. Silva
