

INSTITUTO DE PESQUISA E ESTRATÉGIA ECONÔMICA DO CEARÁ – IPECE
ESTADO DO CEARÁANEXO II – COMUNICADO Nº 012/2026 – DE RESULTADO DA PROVA DISCURSIVA – PRELIMINAR
CONCURSO PARA EMPREGO PÚBLICO N º 001/2026
ESPELHOS DE RESPOSTAS DA PROVA DISCURSIVA

401: ANALISTA DE POLÍTICAS PÚBLICAS – ESTUDOS SOCIOECONÔMICOS

QUESTÃO 1

O governo federal lançou em 2021 um programa de incentivo à conclusão do Ensino Médio para jovens de baixa renda, denominado "Bolsa Estudo". O programa oferece um benefício mensal de R\$ 200 para jovens de 16 a 24 anos que estejam matriculados no Ensino Médio regular e que mantenham frequência escolar mínima de 80%. A implementação ocorreu em dois momentos: no primeiro semestre de 2021, o programa foi implantado em 200 municípios com os menores Índices de Desenvolvimento Humano (IDH) do país (grupo tratado). No segundo semestre de 2022, o programa foi estendido a outros 200 municípios com IDH ligeiramente superior (grupo que servirá como controle na avaliação). A escolha não foi aleatória: os municípios tratados eram significativamente mais pobres e com piores indicadores educacionais.

Você, como analista de políticas públicas do IPEA, precisa estimar o efeito causal do programa sobre a taxa de aprovação no Ensino Médio nos municípios tratados após dois anos de implementação (comparando o biênio 2019- 2020 – pré-tratamento – com o biênio 2023-2024 – pós-tratamento). Você decide utilizar o método de Diferenças em Diferenças (DiD).

Considerando o contexto descrito, explique detalhadamente como você aplicaria o método DiD para estimar o impacto do programa "Bolsa Estudo" sobre a taxa de aprovação no Ensino Médio. Sua resposta deve conter de 15 a 30 linhas e abordar:

- a definição e a lógica do método DiD;
- quais dados seriam necessários e como você organizaria o período pré e pós-tratamento;
- os pressupostos do DiD, com ênfase na hipótese de tendências paralelas e como verificá-la;
- quais as principais ameaças à validade do DiD nesse contexto específico.

ESPELHO DE RESPOSTA

O método Diferenças em Diferenças (DiD) é um quase-experimento utilizado para estimar efeitos causais quando a alocação do tratamento não é aleatória. Sua lógica consiste em comparar a evolução do resultado entre o grupo tratado e o grupo controle antes e depois da intervenção, por meio da expressão: $DiD = (Y_{tratado,pós} - Y_{tratado,pré}) - (Y_{controle,pós} - Y_{controle,pré})$. Esse duplo diferencial elimina vieses de diferenças fixas não observáveis entre os grupos (como características institucionais ou culturais que não mudam no tempo) e captura o impacto líquido do programa, desde que ambos os grupos compartilhem a mesma tendência na ausência do tratamento.

Para aplicar o DiD neste caso, seriam necessários dados municipais da taxa de aprovação no ensino médio, obtidos do Censo Escolar/INEP, para os biênios 2019-2020 (pré-tratamento) e 2023-2024 (pós-tratamento). Também coletaria variáveis de controle como IDH, PIB per capita, gasto público por aluno. Organizaria uma base em painel com dois períodos (pré e pós) e dois grupos: tratado (200 municípios com programa desde 2021) e controle (200 municípios que receberam o programa apenas em 2022, portanto não tratados no período de análise). O marco temporal da intervenção é 2021.

O pressuposto mais crítico do DiD é o de tendências paralelas: na ausência do programa, os grupos tratado e controle teriam evoluído de forma semelhante. Para verificá-lo, comparo a trajetória da taxa de aprovação entre 2019 e 2020 para ambos os grupos – se as médias evoluírem de maneira estatisticamente semelhante (ou com diferença constante), o pressuposto é plausível. Testes de placebo (supondo falsas datas de intervenção) também podem ser realizados. Outros pressupostos incluem ausência de efeitos de transbordamento entre grupos e exogeneidade do tratamento.

As principais ameaças à validade do DiD neste contexto são: violação das tendências paralelas (municípios tratados já melhoravam mais rapidamente antes do programa, superestimando o efeito); efeitos de transbordamento (migração de alunos de municípios controle para tratados, contaminando o controle); eventos concomitantes (políticas estaduais que afetam diferencialmente os grupos); desgaste diferencial (municípios com piores resultados abandonam a coleta de dados); e erro de medição (mudanças nos critérios de aprovação escolar). Para mitigar, recomenda-se incluir controles adicionais e realizar testes de robustez.

QUESTÃO 2

Uma indústria química localizada em um município de médio porte gera, como subproduto de sua produção, emissão de poluentes atmosféricos que afetam a saúde da população residente nas proximidades. Estudos da Secretaria Municipal de Saúde apontam um aumento de 30% nos casos de doenças respiratórias na região vizinha à fábrica nos últimos cinco anos, custando ao sistema público de saúde cerca de R\$ 2 milhões anuais em tratamentos. A indústria, por sua vez, gera empregos e arrecadação de impostos. Atualmente, não há qualquer regulamentação ou tributação específica sobre a poluição emitida. O secretário municipal de meio ambiente contratou você, economista do quadro técnico do município, para elaborar uma proposta de intervenção baseada em fundamentos microeconômicos. Após análise, você conclui que se trata de um caso clássico de externalidade negativa de produção. Com base na teoria microeconômica das externalidades e nos instrumentos de política pública corretivos, explique, utilizando de 15 a 30 linhas e de forma detalhada:

- I. o conceito de externalidade negativa e por que o mercado, sem intervenção, aloca recursos de forma ineficiente nesse caso;
- II. qual instrumento de política pública você recomendaria ao secretário (justificando a escolha) e como ele funcionaria na prática;
- III. quais seriam os efeitos esperados sobre o nível de poluição, o preço do produto industrial, o bem-estar da população e a receita do município;
- IV. uma possível objeção prática à aplicação desse instrumento e como contorná-la.

ESPELHO DE RESPOSTA

A externalidade negativa é um fenômeno em que a produção ou o consumo de um bem impõe custos a terceiros não envolvidos na transação, sem que haja compensação. No caso da indústria química, a poluição gerada afeta a saúde da população vizinha e onera o sistema público de saúde em R\$ 2 milhões anuais, caracterizando um custo externo não refletido no preço do produto industrial. Sem intervenção, o mercado maximiza o lucro privado igualando custo marginal privado (CMP) à receita marginal. Como a empresa ignora o custo marginal externo (CME), o custo marginal social ($CMS = CMP + CME$) é superior ao CMP, levando a uma quantidade produzida maior que a socialmente ótima, resultando em excesso de poluição e perda de bem-estar (peso morto). Diante disso, recomendo a aplicação da Taxa Pigouviana, um imposto corretivo igual ao custo marginal externo, que internaliza a externalidade ao fazer a empresa arcar com o dano causado. Na prática, o município estabelecerá um valor por tonelada de poluente emitido, fiscalizado por medição direta ou estimativa. A alíquota deve ser ajustada para que a nova quantidade produzida iguale o CMS ao benefício marginal social. Os efeitos esperados são: redução do nível de poluição (a empresa investirá em tecnologia limpa ou reduzirá produção), aumento do preço do produto (parte do imposto é repassada ao consumidor, refletindo o custo social), melhora no bem-estar da população (menos doenças respiratórias e menor gasto público em saúde) e geração de receita para o município (que pode ser usada para compensar danos residuais ou reduzir outros impostos). Uma objeção prática relevante é a dificuldade de mensurar com precisão o custo marginal externo (quanto vale cada tonelada de poluente?). Para contornar, pode-se adotar uma abordagem de ajuste gradual: inicia-se com uma alíquota baixa e eleva-se até que a poluição atinja a meta desejada. Métodos de valoração ambiental (preços hedônicos, custo de doenças) também auxiliam na estimativa. Essa solução respeita o princípio do poluidor-pagador e incentiva a inovação sustentável.

402: ANALISTA DE POLÍTICAS PÚBLICAS – ESTUDOS DE GESTÃO PÚBLICA

QUESTÃO 1

O estado de Minas Verde está iniciando um novo ciclo de governo após a posse do governador eleito. A equipe de planejamento precisa elaborar os instrumentos orçamentários exigidos pela Constituição Federal para organizar a ação governamental dos próximos anos. O secretário de planejamento convocou uma reunião técnica para discutir as características, os prazos e as relações entre o Plano Plurianual (PPA), a Lei de Diretrizes Orçamentárias (LDO) e a Lei Orçamentária Anual (LOA). Com base nas disposições da Constituição Federal de 1988, discorra sobre as funções e as características de cada um desses três instrumentos orçamentários, explicando de que forma eles se articulam para orientar o planejamento e a execução orçamentária do estado.

ESPELHO DE RESPOSTA

O candidato deve demonstrar conhecimento das funções e características dos três instrumentos orçamentários previstos no art. 165 da Constituição Federal de 1988, bem como da articulação entre eles.

Quanto ao Plano Plurianual (PPA), o candidato deve explicar que, nos termos do art. 165, § 1º, da Constituição Federal, a lei que o institui estabelecerá, de forma regionalizada, as diretrizes, os objetivos e as metas da administração pública para as despesas de capital e outras delas decorrentes e para as relativas aos programas de duração continuada. O PPA tem vigência de quatro anos, iniciando no segundo ano do mandato do chefe do Executivo e encerrando-se no primeiro ano do mandato subsequente, garantindo continuidade das políticas públicas. É o instrumento de planejamento de médio prazo que confere a direção estratégica da ação governamental.

Quanto à Lei de Diretrizes Orçamentárias (LDO), o candidato deve indicar que, conforme o art. 165, § 2º, da Constituição Federal, ela compreende as metas e prioridades da administração pública, incluindo as despesas de capital para o exercício

financeiro subsequente, orienta a elaboração da LOA, dispõe sobre as alterações na legislação tributária e estabelece a política de aplicação das agências financeiras oficiais de fomento. A LDO funciona como elo entre o PPA e a LOA, selecionando, dentro dos objetivos do PPA, as prioridades para o exercício seguinte.

Quanto à Lei Orçamentária Anual (LOA), o candidato deve apontar que, de acordo com o art. 165, § 5º, da Constituição Federal, ela compreende o orçamento fiscal, o orçamento de investimento das empresas em que o ente detenha a maioria do capital social com direito a voto e o orçamento da seguridade social. A LOA estima as receitas e fixa as despesas para um único exercício financeiro, devendo ser compatível tanto com o PPA quanto com a LDO.

Por fim, o candidato deve explicar a articulação entre os três instrumentos: o PPA define as diretrizes e metas de médio prazo; a LDO seleciona as prioridades do PPA para o exercício seguinte e orienta a elaboração da LOA; e a LOA executa anualmente as ações previstas no PPA segundo as prioridades estabelecidas pela LDO. O art. 167, § 1º, da Constituição Federal veda a inclusão na LOA de dotação para investimento com duração superior a um exercício financeiro que não esteja previsto no PPA ou sem prévia autorização legislativa específica, reforçando a necessidade de compatibilidade entre os instrumentos.

QUESTÃO 2

O município de Serra Dourada encerrou o segundo quadrimestre do exercício financeiro com indícios de que a arrecadação de receitas ficará aquém do previsto na Lei Orçamentária Anual. O secretário de finanças identificou que, se mantida a tendência de queda na arrecadação, as metas de resultado primário estabelecidas para o exercício poderão não ser atingidas. Diante desse cenário, o prefeito solicitou um parecer técnico sobre as medidas que a legislação impõe ao gestor público nessa situação, bem como sobre os estágios da despesa pública que precisam ser observados durante a execução orçamentária. Com base na Lei no 4.320/1964 e na Lei Complementar no 101/2000 (Lei de Responsabilidade Fiscal), disserte sobre os estágios de execução da despesa pública previstos na Lei no 4.320/1964 e sobre o mecanismo de limitação de empenho previsto na Lei de Responsabilidade Fiscal, explicando as condições que o autorizam e seus efeitos sobre a gestão orçamentária e financeira do município.

ESPELHO DE RESPOSTA

O candidato deve demonstrar conhecimento dos estágios de execução da despesa pública previstos na Lei n. 4.320/1964 e do mecanismo de limitação de empenho disciplinado pela Lei de Responsabilidade Fiscal (LC n. 101/2000).

Quanto aos estágios da despesa, o candidato deve indicar que, nos termos dos arts. 58 a 64 da Lei n. 4.320/1964, a execução da despesa orçamentária percorre três estágios: empenho, liquidação e pagamento.

O empenho, previsto no art. 58, é o ato emanado de autoridade competente que cria para o Estado a obrigação de pagamento pendente ou não de implemento de condição; é o ato que reserva o crédito orçamentário para determinada despesa.

A liquidação, disciplinada no art. 63, consiste na verificação do direito adquirido pelo credor, tendo por base os títulos e documentos comprobatórios do respectivo crédito, verificando-se a origem e o objeto do que se deve pagar, a importância exata a pagar e a quem se deve pagar. O pagamento, previsto no art. 64, é o despacho exarado por autoridade competente, determinando que a despesa seja paga; a efetivação do pagamento é feita por tesouraria ou pagadoria regularmente constituídas.

Quanto à limitação de empenho, o candidato deve explicar que, nos termos dos arts. 8º e 9º da LC n. 101/2000, até trinta dias após a publicação dos orçamentos, o Poder Executivo estabelecerá a programação financeira e o cronograma de execução mensal de desembolso. Se verificado, ao final de um bimestre, que a realização da receita poderá não comportar o cumprimento das metas de resultado primário ou nominal estabelecidas no Anexo de Metas Fiscais, os Poderes e o Ministério Público promoverão, por ato próprio e nos montantes necessários, nos trinta dias subsequentes, limitação de empenho e movimentação financeira, segundo os critérios fixados pela lei de diretrizes orçamentárias.

O candidato deve ainda mencionar que, nos termos do art. 9º, § 2º, da LC n. 101/2000, não serão objeto de limitação as despesas que constituam obrigações constitucionais ou legais do ente, inclusive aquelas destinadas ao pagamento do serviço da dívida, as relacionadas a inovação e ao desenvolvimento científico e tecnológico custeadas por fundo criado para tal finalidade e as ressalvadas pela lei de diretrizes orçamentárias.

Por fim, o candidato deve explicar o efeito prático desse mecanismo: ao restringir os limites de empenho, o gestor impede que novas obrigações de gasto sejam assumidas além da capacidade real de arrecadação, preservando o equilíbrio fiscal do exercício e o cumprimento das metas fiscais pactuadas, em consonância com o princípio da responsabilidade fiscal estabelecido no art. 1º da LC n. 101/2000.

403: ANALISTA DE POLÍTICAS PÚBLICAS – GEOPROCESSAMENTO E ESTUDOS TERRITORIAIS

QUESTÃO 1

O Governo do Estado do Ceará pretende implantar novas unidades de atendimento social em municípios com elevados índices de vulnerabilidade socioeconômica. Para subsidiar a tomada de decisão, o IPECE recebeu diferentes bases georreferenciadas contendo informações sobre população, renda, infraestrutura viária, equipamentos públicos existentes e indicadores sociais. Na condição de Analista de Políticas Públicas – Geoprocessamento e Estudos Territoriais, elabore um parecer técnico apresentando como o geoprocessamento pode ser utilizado para identificar áreas prioritárias para

implantação dos novos equipamentos públicos. Em sua resposta, aborde obrigatoriamente:

- as etapas de organização e integração dos dados geoespaciais;
- os procedimentos de análise espacial a serem adotados;
- os critérios para definição das áreas prioritárias;
- os produtos cartográficos e indicadores que poderiam subsidiar a tomada de decisão governamental.

ESPELHO DE RESPOSTA

A utilização do geoprocessamento é fundamental para subsidiar a implantação de novos equipamentos públicos em áreas com maior vulnerabilidade socioeconômica. Inicialmente, devem ser integradas em ambiente SIG bases georreferenciadas contendo informações demográficas, renda, infraestrutura viária, equipamentos públicos existentes e indicadores sociais. Após a organização e padronização dos dados, podem ser aplicadas técnicas de análise espacial, como sobreposição de camadas temáticas, análise de proximidade (buffer), análise de acessibilidade e análise multicritério. Essas técnicas permitem identificar áreas com elevada concentração populacional, baixa cobertura de serviços públicos e maior vulnerabilidade social. A definição das áreas prioritárias deve considerar critérios como densidade populacional, distância dos equipamentos existentes, acessibilidade da população, indicadores de renda e índices de vulnerabilidade social. A ponderação dessas variáveis possibilita a construção de um índice territorial de prioridade para investimentos. Como produtos finais, podem ser elaborados mapas temáticos, mapas de áreas prioritárias, painéis geoespaciais e indicadores territoriais que permitam comparar diferentes regiões. Esses resultados contribuem para a tomada de decisão baseada em evidências, tornando mais eficiente a alocação dos recursos públicos e ampliando a efetividade das políticas sociais.

QUESTÃO 2

Após um período de chuvas intensas, diversos municípios cearenses registraram ocorrências de inundações urbanas, deslizamentos de encostas e danos à infraestrutura pública. O Governo do Estado solicitou ao IPECE a elaboração de um estudo técnico para identificar áreas prioritárias para ações preventivas e mitigação de riscos socioambientais. Na condição de Analista de Políticas Públicas – Geoprocessamento e Estudos Territoriais, apresente uma proposta metodológica para realização desse estudo. Em sua resposta, aborde obrigatoriamente:

- os conceitos de ameaça, vulnerabilidade e risco;
- as informações territoriais necessárias para a análise;
- os procedimentos para identificação e mapeamento das áreas críticas;
- a utilização dos resultados na formulação, monitoramento e avaliação de políticas públicas de prevenção de desastres.

ESPELHO DE RESPOSTA

A análise de riscos socioambientais deve partir da compreensão de que o risco resulta da interação entre a ameaça e a vulnerabilidade dos elementos expostos. As ameaças correspondem aos eventos potencialmente danosos, como inundações e deslizamentos, enquanto a vulnerabilidade está relacionada às condições sociais, econômicas e ambientais da população afetada. Para a realização do estudo, devem ser utilizadas informações territoriais referentes ao relevo, rede de drenagem, uso e ocupação da terra, cobertura vegetal, densidade populacional, infraestrutura urbana e histórico de ocorrências de desastres. Em ambiente SIG, essas informações podem ser integradas por meio de técnicas de geoprocessamento, permitindo a elaboração de mapas de suscetibilidade, vulnerabilidade e risco. A análise espacial possibilita identificar áreas críticas e definir prioridades para intervenções preventivas. Os resultados podem subsidiar ações de planejamento urbano, obras de contenção, sistemas de alerta, reassentamento de populações vulneráveis e monitoramento permanente das áreas de risco. Dessa forma, o estudo contribui para a formulação, implementação e avaliação de políticas públicas voltadas à redução de desastres e ao aumento da resiliência territorial.

404: ANALISTA DE POLÍTICAS PÚBLICAS – MÉTODOS QUANTITATIVOS

QUESTÃO 1

O IPECE foi incumbido de elaborar um diagnóstico territorial da pobreza no Ceará usando microdados da PNAD Contínua e dados do Censo Demográfico. A equipe planeja calcular indicadores demográficos relevantes (razão de dependência, taxa de mortalidade infantil e índice de envelhecimento) por macrorregião. Explique a interpretação da razão de dependência e do índice de envelhecimento no contexto de formulação de políticas sociais para o Ceará.

ESPELHO DE RESPOSTA

- Razão de Dependência Total: Calculada como $[(Pop. 0-14) + (Pop. 65+)] / (Pop. 15-64) \times 100$. Mede quantas pessoas economicamente dependentes (crianças e idosos) cada 100 pessoas em idade ativa (15-64 anos) precisam sustentar. Macrorregiões com alta dependência jovem (comum em sertões) demandam políticas de saúde materno-infantil, educação e creches. Já regiões com alta dependência idosa (cidades com envelhecimento acelerado) exigem foco em previdência social, saúde do idoso e instituições de longa permanência.

- Índice de Envelhecimento: Calculado como $(Pop. 65+/Pop. 0-14) \times 100$. Indica a proporção de idosos em relação a crianças e jovens, refletindo a inversão da pirâmide etária. Um índice crescente, como observado no Ceará, sinaliza a necessidade de adaptar os serviços públicos para as demandas do envelhecimento populacional, como mobilidade urbana

acessível, saúde para doenças crônicas e sistemas de seguridade social robustos.

QUESTÃO 2

Uma equipe do IPECE avalia o impacto de um programa de capacitação profissional para jovens de 18 a 29 anos de municípios cearenses, implementado em 2019. A base de dados contém informações de 2016 a 2023 sobre: renda formal (R\$ mensais), tempo até primeiro emprego formal após o programa (em meses) e município de residência. O programa foi oferecido apenas aos municípios com IDM abaixo da mediana do Ceará. Com base nessas informações, proponha um método de avaliação de impacto adequado para estimar o efeito causal do programa sobre a renda formal, justificando a escolha e discutindo os pressupostos necessários.

ESPELHO DE RESPOSTA

O desenho do programa, que foi implementado apenas em municípios com IDM abaixo da mediana do Ceará, sugere a aplicação de dois métodos de avaliação de impacto robustos:

1 Regressão Descontínua (RDD):

- Justificativa: O RDD é adequado devido à existência de um ponto de corte claro (a mediana do IDM) que determina a elegibilidade para o programa. Municípios marginalmente abaixo (tratados) e acima (controles) da mediana são considerados comparáveis, permitindo a identificação do Efeito Médio do Tratamento Local (LATE) no ponto de corte.
- Pressupostos: O principal pressuposto é que os municípios não podem manipular seu IDM para cruzar o ponto de corte. Isso pode ser testado empiricamente através do teste de McCrary, que verifica a continuidade da densidade das unidades de tratamento em torno do ponto de corte.

2 Diferenças-em-Diferenças (DiD):

- Justificativa: Com dados em painel (2016–2023) e variação temporal (antes de 2019 = pré-programa; após = pós-programa) e espacial (municípios que receberam vs. não receberam o programa), o DiD é uma alternativa viável. Ele compara a mudança na renda formal dos jovens nos municípios tratados com a mudança nos municípios de controle ao longo do tempo.
- Pressupostos: O pressuposto central é o das tendências paralelas: na ausência do programa, a renda dos jovens nos municípios tratados teria seguido a mesma trajetória (tendência) que a renda dos jovens nos municípios de controle. Este pressuposto pode ser testado visualmente e estatisticamente através de um event study no período pré-tratamento (2016–2018), verificando se os coeficientes de interação ano $\times$ tratamento são não significativos antes de 2019.

405: ANALISTA DE POLÍTICAS PÚBLICAS – COMPUTAÇÃO E CIÊNCIA DE DADOS

QUESTÃO 1

O Ministério do Desenvolvimento Social pretende desenvolver uma plataforma denominada "AlvoSocial" para apoiar a identificação de famílias em situação de insegurança alimentar severa. Para isso, será utilizado um modelo de Aprendizado de Máquina Supervisionado treinado com dados históricos de programas sociais, permitindo estimar a probabilidade de vulnerabilidade socioeconômica das famílias cadastradas. Considerando o cenário apresentado e os fundamentos de Aprendizado de Máquina, elabore um texto dissertativo abordando, obrigatoriamente, os seguintes aspectos:

- explique o conceito de Aprendizado de Máquina Supervisionado e sua aplicação no problema descrito;
- indique e justifique qual métrica de avaliação deve ser priorizada para esse modelo preditivo, considerando que o objetivo prioritário da política pública é mitigar o risco de exclusão de famílias vulneráveis;
- explique o que são Falsos Positivos e Falsos Negativos no contexto apresentado e discuta qual deles representa o maior risco para a política pública, sob a premissa de que a omissão do auxílio é o cenário mais grave;
- discorra sobre o problema do viés algorítmico (algorithmic bias), explicando como dados históricos podem influenciar as decisões futuras do modelo e gerar efeitos discriminatórios ou de exclusão social.

ESPELHO DE RESPOSTA

1. Aprendizado de Máquina Supervisionado

O candidato deve demonstrar compreensão de que o Aprendizado de Máquina Supervisionado consiste no treinamento de modelos a partir de dados históricos previamente rotulados, permitindo que o algoritmo aprenda padrões e realize previsões sobre novos registros.

No cenário proposto, o modelo utiliza dados históricos de programas sociais para identificar famílias com maior probabilidade de se encontrarem em situação de insegurança alimentar severa.

2. Métricas de Avaliação e Impacto Social

O candidato deve identificar adequadamente a métrica Sensibilidade (Recall) como a mais apropriada para o problema proposto. Também será aceita a indicação de outras métricas focadas em cobertura ou adequadas para cenários desbalanceados, tais como variações do F-beta Score (como o F2-Score ou valores de $\beta > 1$), PR-AUC (Precision-Recall AUC) ou matrizes de custo customizadas, desde que o preenchimento de requisitos técnicos seja devidamente justificado.

No contexto da plataforma AlvoSocial, espera-se a conceituação correta dos erros de classificação:

• Falso Positivo (FP): Ocorre quando o modelo classifica uma família como em situação de insegurança alimentar severa, quando na realidade ela não está.

• Falso Negativo (FN): Ocorre quando o modelo classifica uma família como não vulnerável, quando na realidade ela se encontra em situação de insegurança alimentar severa.

A justificativa técnica esperada é que, em políticas públicas voltadas à assistência social, o erro mais grave é o Falso Negativo, situação em que uma família vulnerável deixa de receber auxílio governamental. O candidato deve demonstrar entendimento de que um Falso Positivo gera custo financeiro adicional para a Administração Pública (atendimento de quem não precisava), enquanto um Falso Negativo compromete diretamente a subsistência de famílias em risco, configurando um custo social muito superior.

3. Viés Algorítmico (Algorithmic Bias)

O candidato deve explicar que modelos de Aprendizado de Máquina reproduzem padrões existentes nos dados utilizados para treinamento.

Podem ser citados exemplos como sub-representação de determinados grupos populacionais, falhas cadastrais, exclusão geográfica, desigualdades socioeconômicas históricas, vieses institucionais, ou ainda fenômenos estritamente estatísticos como desbalanceamento de classes, erro de amostragem, viés de seleção/medição e concept drift histórico.

Qualquer abordagem que conecte a falha ou limitação dos dados históricos com a perpetuação de injustiças, distorções ou erros de exclusão sistemática em previsões futuras atende ao quesito, demonstrando como o viés compromete a efetividade das políticas públicas.

QUESTÃO 2

A Controladoria-Geral de um estado da Federação pretende aprimorar seus mecanismos de auditoria preventiva em contratações públicas. Para isso, foi desenvolvido um ambiente analítico contendo informações sobre milhares de licitações realizadas pelos órgãos estaduais, incluindo características dos objetos contratados, quantidades adquiridas, preços praticados e demais atributos relacionados às compras públicas. Como não existem registros históricos suficientemente confiáveis que indiquem previamente quais processos apresentaram superfaturamento, cartelização ou direcionamento de edital, a equipe de Ciência de Dados decidiu empregar técnicas de mineração de dados para analisar automaticamente processos licitatórios com características semelhantes, permitindo posteriormente a identificação de possíveis irregularidades. Considerando o cenário apresentado e os fundamentos de Análise de Padrões e Mineração de Dados (Data Mining), elabore um texto dissertativo abordando, obrigatoriamente, os seguintes aspectos:

1. identifique a tarefa de mineração de dados e o paradigma de aprendizado de máquina mais adequados para o problema apresentado, justificando sua escolha com base nas características do cenário;
2. explique como técnicas de agrupamento (clustering) ou algoritmos especializados em detecção de anomalias podem auxiliar na identificação de padrões de comportamento e na detecção de possíveis irregularidades em processos licitatórios;
3. explique por que a Distância Euclidiana pode se tornar inadequada em bases de dados com grande quantidade de atributos, abordando o fenômeno conhecido como maldição da dimensionalidade (curse of dimensionality) e seus impactos na análise dos dados.

ESPELHO DE RESPOSTA

1. Paradigma de Aprendizado e Tarefa de Mineração de Dados:

O candidato deve reconhecer que o cenário proposto exige o paradigma de Aprendizado Não Supervisionado, justificando obrigatoriamente sua escolha com base na inexistência de dados históricos rotulados (ou seja, a falta de marcação prévia sobre quais licitações foram fraudulentas, cartelizadas ou superfaturadas).

Quanto à tarefa de mineração de dados, o candidato deve indicar adequadamente o Agrupamento (Clustering) ou a Detecção de Anomalias (Outlier Detection). Podem ser citados algoritmos tradicionais de agrupamento (como K-Means, DBSCAN ou métodos hierárquicos) ou algoritmos nativos de isolamento/anomalia (como Isolation Forest, Local Outlier Factor - LOF, ou One-Class SVM), explicando que o objetivo é extrair estruturas intrínsecas e padrões ocultos diretamente dos atributos das compras públicas.

2. Identificação de Padrões e Detecção de Anomalias:

O candidato deve demonstrar o entendimento prático de como essas técnicas funcionam no contexto de auditoria de licitações:

• Abordagem por Agrupamento: Os algoritmos dividem as licitações em grupos homogêneos (clusters) com base na similaridade de seus atributos (ex: quantidade, objeto, preço). Uma vez formados os grupos normais, os registros que ficam muito distantes dos centros dos grupos, ou que são isolados em grupos minúsculos, são identificados como desvios ou potenciais irregularidades.

• Abordagem por Detecção de Anomalias Nativa: Os algoritmos tentam isolar ou identificar diretamente os pontos que violam o comportamento estatístico predominante da base, apontando os processos licitatórios discrepantes.

Em ambos os casos, deve ficar claro que a técnica estatística não carimba ou decreta a existência final de fraude, mas atua como um filtro analítico e preventivo para que os auditores humanos concentrem seus esforços fiscalizatórios nos casos suspeitos.

3. Maldição da Dimensionalidade e a Distância Euclidiana:

O candidato deve abordar o fenômeno da Maldição da Dimensionalidade (Curse of Dimensionality) em bases de alta dimensionalidade (grande volume de atributos). Deve ser explicado que, à medida que o número de dimensões aumenta, o volume do espaço cresce exponencialmente, tornando os dados altamente esparsos.

Nesse cenário geométrico, as distâncias (como a Distância Euclidiana) entre quaisquer dois pontos da base tendem a convergir para o mesmo valor, tornando-se indistinguíveis. Como consequência direta, perde-se o poder discriminatório da métrica de distância, o que prejudica severamente a capacidade dos algoritmos de definir vizinhanças significativas e formar agrupamentos ou isolar anomalias com qualidade.