



Universidade de São Paulo



CONCURSO ANALISTA DE SISTEMAS (ESPECIALIDADE: CIÊNCIA DE DADOS) PARA A UNIVERSIDADE DE SÃO PAULO  
EDITAL RH Nº 13/2023

Instruções

1. **Só abra este caderno quando o fiscal autorizar.**
2. Verifique se o seu nome está correto na capa deste caderno e se a folha de respostas pertence ao **grupo J**. Informe ao fiscal de sala eventuais divergências.
3. Durante a prova, são **vedadas** a comunicação entre candidatos e a utilização de qualquer material de consulta e de aparelhos de telecomunicação.
4. Duração da prova: **05 (cinco) horas**. Cabe ao candidato controlar o tempo a partir do relógio disponibilizado na sala de provas. O(A) candidato(a) poderá retirar-se da sala definitivamente somente após decorridas **02 (duas) horas** de prova. Não haverá tempo adicional para preenchimento da folha de respostas.
5. Lembre-se de que a FUVEST se reserva o direito de efetuar procedimentos adicionais de identificação e controle do processo, visando a garantir a plena integridade do exame. Assim, durante a realização da prova, poderá ser coletada por um fiscal uma **foto** do(a) candidato(a) para fins de reconhecimento facial, para uso exclusivo da USP e da FUVEST. A imagem não será divulgada nem utilizada para quaisquer outras finalidades, nos termos da lei.
6. Após a autorização do fiscal da sala, verifique se o caderno está completo. Ele deve conter **60 (sessenta)** questões objetivas, com 05 (cinco) alternativas cada, das quais apenas uma atende ao enunciado, e **02 (duas)** questões dissertativas. Informe ao fiscal de sala eventuais divergências quanto ao número de questões e de alternativas.
7. Preencha as folhas de respostas com cuidado, utilizando caneta esferográfica de **tinta azul ou preta**. As folhas de respostas **não serão substituídas** em caso de rasura.
8. Ao final da prova, é **obrigatória** a devolução das folhas de respostas acompanhadas deste caderno de questões.

**Declaração**

Declaro que li e estou ciente das informações que constam na capa desta prova, nas folhas de respostas, bem como nos avisos que foram transmitidos pelo fiscal de sala.

\_\_\_\_\_  
ASSINATURA

O(a) candidato(a) que não assinar esta capa será considerado(a) ausente da prova.



RASCUNHO



## TEXTO PARA AS QUESTÕES 01 A 03

## O jornalismo além da objetividade

Para a *exígua* parcela do improvável leitorado que ainda se interessa pelos estudos de jornalismo, acaba de sair um documento de leitura obrigatória: *Além da objetividade – produzindo noticiário confiável nas redações atuais*. [...]

Os dois autores são nomes consagrados na profissão. Leonard Downie Jr., professor da Escola Walter Cronkite, fez carreira no *Washington Post*, onde chegou a editor executivo. Andrew Heyward, docente na mesma faculdade, foi presidente da *CBS News* entre 1996 e 2005. Depois de consultarem uma respeitável bibliografia, a dupla entrevistou 76 pessoas que exercem cargos-chave na imprensa dos Estados Unidos e chegou a uma conclusão nada trivial: a palavra “objetividade”, tão cara à tradição dos jornais, anda fora de moda (*outmoded*). Repórteres e editores não têm mais o mesmo gosto em pronunciá-la. O termo já não nomeia o requisito central da credibilidade, pois “perdeu seu poder de definir os padrões mais altos da excelência jornalística”. [...]

Não se trata de negligenciar a realidade, de jeito nenhum, mas de olhar mais longe. Trata-se de examinar o pano de fundo e de decifrar as opiniões fundamentadas que entram em conflito. O texto jornalístico só é bom de verdade quando, além de narrar o acontecido, transpira pensamento. Só assim vai *refletir o real e refletir sobre o real*.

Para resumir, o que entrou em crise não é a tentativa de captar os dados objetivos da realidade, mas a *empáfia* com que muitos desfraldavam a bandeira da objetividade. Não dá para continuar assim. Já não tem serventia o repórter que descreve olímpicamente um episódio qualquer, ouve uma fonte favorável e outra contrária e, com isso, dá o trabalho por encerrado – o cidadão que se vire para encontrar a conclusão. A imprensa responsável não tem parte com a indiferença. Ou ela vibra junto com a audiência ou ficará isolada.

É nesse sentido que a boa redação jornalística procura desvelar as forças que se batem para fazer prevalecer uma interpretação ou outra, deixa claro seu método de trabalho, abre os olhos para a diversidade e compartilha com o público os valores e princípios que a orientam. Tudo se resume a uma questão de honestidade, em três frentes simultâneas: honestidade para relatar o que aconteceu, para esmiuçar o contexto e, em terceiro lugar, para não esconder seus próprios compromissos.

O ideal da precisão fria – que sempre foi uma forma de impostura positivista – caducou. Acima dele deve estar a relação franca com a audiência. O jornalismo digno de confiança respeita as expectativas de veracidade das suas fontes e de seus públicos, igualmente, do mesmo modo que respeita sua coerência interna. Assim, tece o diálogo entre sujeitos ativos num padrão civilizado e pacífico. Em outras palavras, o jornalismo se faz na intersubjetividade racional. [...]

Fonte: BUCCI, Eugênio. O jornalismo além da objetividade. Estadão (online), 09 de março de 2023 (adaptado).

## 01

De acordo com o texto, é correto afirmar que

- (A) o ideal de objetividade, uma impostura positivista, foi responsável pela crise de credibilidade do jornalismo.
- (B) a intersubjetividade deve superar a objetividade no jornalismo contemporâneo, conforme se faz tradicionalmente nos Estados Unidos.
- (C) a credibilidade de um jornal se constrói pela clareza quanto aos seus compromissos e pela honestidade no relato e na contextualização.
- (D) a boa redação jornalística deve privilegiar a conexão com os valores e os princípios da audiência em detrimento da factualidade.
- (E) a objetividade não é mais central ao jornalismo, pois o leitor contemporâneo perdeu a capacidade de encontrar a conclusão por conta própria.

## 02

A construção “refletir o real e refletir sobre o real” (terceiro parágrafo) refere-se à necessidade de o bom jornalismo

- (A) manter-se fiel aos fatos, isento, deixando às fontes o papel de refletir sobre os acontecimentos.
- (B) escolher o ponto de vista que julga mais pertinente em relação a seus compromissos, tratando-o como o real.
- (C) vibrar com a audiência, espelhando o que esse grupo considera mais relevante e narrando os acontecimentos a partir dessa perspectiva.
- (D) ancorar-se nos acontecimentos concretos e na diversidade de posições a fim de produzir interpretações embasadas para o leitor.
- (E) assumir que a realidade é, no fundo, uma construção intersubjetiva que prescinde de racionalidade.

## 03

As palavras “exígua” e “empáfia” (primeiro e quarto parágrafos) podem ser substituídas, sem prejuízo de sentido no texto, por:

- (A) Expressiva e coragem.
- (B) Considerável e elegância.
- (C) Inexorável e vaidade.
- (D) Singela e insensatez.
- (E) Diminuta e soberba.



## TEXTO PARA AS QUESTÕES 04 E 05

Ataque em escola expõe carência de ensino inclusivo a crianças e adolescentes

*Lei federal incentiva combate ao assédio moral sistemático, mas instituições precisam adotar medidas concretas contra o preconceito, a exclusão e a violência.*

O recente ataque de um aluno de 13 anos em uma escola estadual de São Paulo, que resultou na morte de uma professora e cinco pessoas feridas, não é um caso isolado.

Ele retrata comportamentos moldados pela mais absoluta carência de valores e indiferença a sentimentos alheios, somados à ineficácia de gestões públicas na implementação de práticas inclusivas, formação continuada, valorização de professores e aplicação de leis.

Um grande erro é pensar que o bullying e o preconceito estão atrelados a uma mera vocação violenta.

Essa visão nos faz buscar vilões e heróis na história, acreditando que basta retirarmos os vilões que o problema será extirpado.

O aumento da violência dentro das escolas está diretamente ligado à ausência de práticas inclusivas e preventivas nas instituições de ensino. Ignorar isso é um equívoco perigoso.

O respeito à diversidade precisa ser ensinado nas escolas, locais repletos de pessoas em formação e distintas umas das outras, um mundo bem diferente do ambiente doméstico.

Pesquisa feita pela FEBRABAN-IPESPE mostrou a escola como ambiente mais propício a esse tipo de assédio moral — sendo cor, raça e orientação sexual as principais razões para o bullying. O estudo "Bullying e Cancelamento: Impacto na Vida do Brasileiro" ouviu 3.000 pessoas nas cinco regiões do país em 2022.

A violência precisa ser tratada com seriedade por toda a comunidade escolar. É preciso que as instituições de ensino, tanto públicas quanto privadas, adotem medidas efetivas de combate e prevenção ao bullying, conforme estabelece a lei 13.185/2015 – Combate à Intimidação Sistemática (Bullying).

A lei informa, por exemplo, que é dever do estabelecimento de ensino assegurar medidas de conscientização, prevenção, diagnose e combate à violência e à intimidação sistemática.

A educação inclusiva traz resultados efetivos em uma escola quando desenvolvida com uma rede de profissionais especializados em educação inclusiva. O núcleo deve agir de forma abrangente nas comunidades escolares, envolvendo todas as questões de diversidade e preconceitos.

*Esse grupo especializado se movimenta usando todos os seus tentáculos.* Atua diretamente nas formações continuadas de professores, das famílias e de alunos para que todos se tornem agentes de mudança e contribuam para um ambiente respeitoso e que gere muito mais garantia de aprendizagem. [...]

Fonte: VIDEIRA, Carolina. Ataque em escola expõe carência do ensino inclusivo a crianças e adolescentes. Folha UOL (online), 30 mar. 2023 (adaptado).

## 04

Segundo o texto, a violência na escola

- (A) pode ser combatida por ações formativas que envolvem professores, famílias e alunos.
- (B) decorre da ausência de legislação pertinente sobre o tema.
- (C) é resultado da diversidade de valores das pessoas em formação e que são distintas umas das outras.
- (D) é efeito de projetos mal conduzidos de educação inclusiva, que estimulam práticas de assédio moral.
- (E) não é um problema sistêmico, podendo ser combatida com ações isoladas em cada comunidade.

## 05

Em “Esse grupo especializado se movimenta usando todos os seus tentáculos” (último parágrafo), observa-se o uso de

- (A) metáfora.
- (B) personificação.
- (C) hipérbole.
- (D) sinestesia.
- (E) eufemismo.

## 06

Leia o texto de campanha apresentado.

Dia Mundial de luta contra a AIDS  
Prevenir é sempre a melhor escolha  
Use camisinha e, na dúvida, faça o teste.  
Com o diagnóstico precoce e o tratamento adequado, o vírus HIV torna-se indetectável.  
Procure uma unidade de saúde.

Fonte: UniEduK. Ação da Campanha de Prevenção ao HIV/AIDS acontece dia 09/12 da UniFAJ. UniEduk (online), Notícias, 07 dez. 2021.

Em “Com o diagnóstico precoce e o tratamento adequado, o vírus torna-se indetectável” (em destaque), depreende-se uma relação lógico-semântica de

- (A) alternatividade.
- (B) concessão.
- (C) condição.
- (D) finalidade.
- (E) proporção.



## TEXTO PARA AS QUESTÕES 07 A 10

“Quiet Quitting”, fenômeno nas redes sociais, é uma forma de reação à vida real

O *Quiet Quitting* se tornou conhecido após ganhar as redes sociais, mais especificamente o TikTok, nas quais diversos perfis compartilharam o que seria esse fenômeno e como aderir ao movimento. Em 2020, os Estados Unidos se viram frente a um movimento que ganhou o nome de *A Grande Renúncia*, o qual reverbera até hoje e levou 4,5 milhões de americanos à demissão voluntária só no mês de maio.

O *Quiet Quitting* está, de certa forma, relacionado a essa renúncia em massa. “É um termo que, em tradução livre, quer dizer ‘demissão silenciosa’. E ele diz respeito ao comportamento de fazer o mínimo no trabalho”, explica Natália Lins Brandão, pesquisadora do Instituto de Psicologia da USP. Uma das causas pode ser que, no período da pandemia, as fronteiras entre horário de trabalho e horário de lazer, assim como o próprio estado físico da casa e do trabalho, acabaram se misturando. Isso levou à completa exaustão, pois a preocupação virou um trabalho de 24h por dia. Nesse período, muitas pessoas viram que seu trabalho poderia ser feito remotamente, sem perder a produtividade.

Superficialmente, diz-se que é um movimento geracional que tem a ver com a falta de querer ou a *desmotivação* para trabalhar. *Isso faria com que muitos trabalhadores desistissem de seus empregos ou não cumprissem com mais do que o combinado na hora da contratação.* Na contramão, a demissão silenciosa é muito mais que isso. Não se trata, assim, apenas de um desânimo ou de quem opta por fazer o mínimo, mas pode ser uma resposta à cobrança excessiva de produtividade e entrega. Muitos não veem futuro na empresa em que estão empregados, estão psicologicamente separados de seu trabalho ou não satisfeitos com a descrição do cargo. Também, a maioria das pessoas que começam a agir dessa forma está procurando por novos empregos. [...]

## O papel dos gestores

De acordo com um estudo publicado no *Harvard Business Review*, o *Quiet Quitting* “tem mais a ver com a inabilidade dos gestores de manterem uma boa comunicação do que propriamente com a falta de vontade dos empregados. Confiar na sua liderança influencia muito em como se portar no trabalho e, quanto mais um líder abertamente conversa com seu subordinado, maior é o nível de confiança. Isso resulta em um sentimento de que seu trabalho tem algum propósito, que o esforço vale a pena e que o gestor se importa com seu bem-estar”.

Natália, porém, lembra que esse fenômeno não atinge a classe trabalhadora por inteiro: “*Isso não é hegemônico*, tem um recorte de classe”. A pesquisadora ainda salienta que pessoas que não podem escolher entre trabalhar ou não, muitas vezes não podem optar pelo *Quiet Quitting*.

Fonte: ESTANISLAU, Julia. “Quiet Quitting”, fenômeno nas redes sociais, é uma forma de reação à vida real. *Jornal da USP* (online), 01 nov. 2022 (adaptado).

## 07

De acordo com o texto, o *quiet quitting*

- (A) é típico de uma geração que não tem vontade de trabalhar, nunca está satisfeita com o trabalho e não sabe se relacionar com os gestores.
- (B) é um comportamento que atinge toda uma geração independentemente da classe social.
- (C) é um fenômeno que ganhou força na pandemia, momento em que as fronteiras entre o pessoal e o profissional se diluíram.
- (D) é um movimento que conquistou as redes sociais, especialmente o *TikTok*, em resposta a decisões de gestores de demitir em massa trabalhadores durante a pandemia.
- (E) é uma proposta de reconfiguração do ambiente de trabalho às exigências das novas gerações, que desejam uma progressão de carreira mais rápida.

## 08

Em “Isso não é hegemônico” (quinto parágrafo), o pronome “isso” refere-se

- (A) ao *quiet quitting*.
- (B) à classe trabalhadora.
- (C) ao estudo publicado no *Harvard Business Review*.
- (D) à confiança nos líderes.
- (E) à cobrança excessiva de produtividade.

## 09

Em “Isso faria com que muitos trabalhadores desistissem de seus empregos ou não cumprissem com mais do que o combinado na hora da contratação” (terceiro parágrafo), o verbo “fazer” está conjugado no Futuro do Pretérito. Caso o conjuguemos no Futuro do Presente, as três formas verbais grifadas, de acordo com a norma culta da língua portuguesa, passariam a:

- (A) Fará, desistirão, cumprirão.
- (B) Faz, desistam, cumpram.
- (C) Faz, desistirem, cumprirem.
- (D) Fará, desistam, cumpram.
- (E) Faz, desistirem, cumprirem.

## 10

Assinale a alternativa em que se encontra uma palavra formada pelo(s) mesmo(s) processo(s) que “desmotivação” (terceiro parágrafo):

- (A) Produtividade.
- (B) Estudo.
- (C) Bem-estar.
- (D) Desânimo.
- (E) Inabilidade.



## TEXTO PARA AS QUESTÕES 11 A 13

'A Spiraling Loop of Feedbacks':  
Worst-Case Scenario for Amazon Rainforest

A paper to be published in the *Journal Science* on January 27 has found that humans have degraded more than one-third of the remaining trees in the Amazon rainforest. This degradation could eventually lead to "a spiraling loop of feedbacks," Jos Barlow, a professor of conservation science at Lancaster University in the U.K. and co-author of the paper, told Newsweek.

Up to 38 percent of the remaining Amazon has been affected by human actions, researchers from Brazil's University of Campinas (Unicamp), the Amazon Environmental Research Institute (IPAM), National Institute for Space Research (INPE), and Lancaster University found.

The degradation of this area—equivalent to 5.5 times the size of the state of California—releases carbon emissions equivalent to or greater than those from deforestation.

The Amazon contributes 16 percent of all the land-based photosynthesis in the world, and strongly regulates global carbon and water cycles, sucking in carbon dioxide and producing oxygen. Additionally, *despite* only covering around 0.5 percent of the Earth's surface, the Amazon is home to over 10 percent of all named plant and vertebrate species on Earth.

"Healthy rainforests provide amazing habitat for biodiversity—this is what the Amazon is most famous for," Sally Thompson, an ecohydrologist at The University of Western Australia, told Newsweek. "They usually support clean water in rivers, make it rain, and cool the surrounding area. You can hunt, harvest timber or foods sustainably from healthy and well-managed forests. And a healthy forest can often recover from disturbance. Degraded forests aren't as good at doing any of those things, and often they struggle to recover from disturbance."

Deforestation involves a loss of the forest canopy and a change in land use (e.g., from forest to agriculture or urban land use), while degradation is a process affecting the remaining forests. Degradation essentially means that there is still forest in place but it is not as healthy or as good at providing benefits for the environment or for people.

THOMSON, Jess. 'A Spiraling Loop of Feedbacks': Worst-Case Scenario for Amazon Rainforest. Newsweek (online), 26 jan. 2023 (adaptado).

**11**

De acordo com o texto, a degradação da Floresta Amazônica tem como decorrência

- (A) a diminuição de um terço da mata restante na região.
- (B) a redução em 38 por cento de floresta afetada.
- (C) o aumento de protestos organizados por pesquisadores do meio ambiente e ecologistas.
- (D) a emissão de carbono igual ou superior àquela causada por desmatamento.
- (E) o crescimento proporcional de áreas desmatadas.

**12**

De acordo com o texto, a degradação e o desmatamento são processos diferentes, pois

- (A) referem-se a quantidades diversas de área florestal.
- (B) o desmatamento afeta árvores remanescentes.
- (C) são estudadas por pesquisadores de áreas diversas do campo científico.
- (D) as regiões devastadas deixam de beneficiar ora o meio ambiente, ora as pessoas.
- (E) a degradação envolve a perda de saúde da mata existente.

**13**

Leia a sentença a seguir:

"Up to 38 percent of the remaining Amazon has been affected by human actions, researchers from Brazil's University of Campinas (Unicamp), the Amazon Environmental Research Institute (IPAM), National Institute for Space Research (INPE), and Lancaster University found."

Assinale a alternativa que apresenta sentença cujo uso da expressão "up to" é semelhante ao empregado no trecho apresentado.

- (A) They can really do the task, but they do not seem *up to* it.
- (B) Usually, levels range *up to* intermediate here.
- (C) It is *up to* the jury to decide on the verdict.
- (D) Being online for so long, we thought they were *up to* something.
- (E) Data suggests that the documents have been kept *up to* date according to scientific research.

## TEXTO PARA AS QUESTÕES 14 A 16

Generic catastrophic poverty when selfish investors exploit a  
degradable common resource

Game theory deals with situations in which a number of agents compete with each other, with each participant trying to maximize his or her own profit individually. One speaks of a "Nash equilibrium" if players cannot increase their returns further. The "Tragedy of the Commons" is a game theoretical scenario in which the actors do not compete directly, but indirectly: If someone takes a piece of a common pie, there will be less for everybody else.

Instead of investigating how to avoid the "Tragedy of the Commons," Claudius Gros from Goethe University's Institute for Theoretical Physics examined the resulting Nash equilibrium, with unexpected results: If a common good is divided more or less equally among N interested parties, then *each* receives a share of the order 1/N. However, the respective investment costs still need to be deducted.

Gros' calculations show that, in equilibrium, the actors increase their engagement until the resulting investment costs almost reach the value of the resources the individual investor can secure for her- or himself. Mathematically, the theoretical



physicist was able to show that the final profit of the individual investor scales as  $1/N^2$ .

The original expectation, that investors each receive a proportional share from the resource, remains correct, as Gros' research shows. However, this does not translate into an overall return of the same proportion, which is smaller by a power in the number of investors. Gros denotes the dramatic deterioration of the net profit as "catastrophic poverty," as it implies that unregulated competition drives the individual actor close to the profitability limit, viz to the subsistence level.

Similarly, Gros was able to show that catastrophic poverty can be avoided when the actors cooperate with each other. Cooperation leads to a net profit corresponding to the number of investors in simple power, the classical result.

The result of the investigations is therefore that the "Tragedy of the Commons" can cause substantially more damage than previously assumed. Uncontrolled access not only leads to a potentially excessive exploitation of the resource, a topic that has been the focus of many previous studies. In addition, investors suffer themselves when only maximizing their own profits.

GROS, Claudius. Generic catastrophic poverty when selfish investors exploit a degradable common resource. Royal Society Open Science (online), 08 Feb. 2023 (adaptado)

## 14

De acordo com o texto, a "Tragédia dos Comuns" pode ser definida como

- (A) as questões éticas envolvidas na tentativa de maximizar os lucros.
- (B) os danos psicológicos causados em investidores de alto risco.
- (C) uma competição indireta em que o consumo de um bem comum por um participante leva à menor quantidade desse bem para todos.
- (D) uma divergência entre diferentes partes interessadas por causa de um bem comum.
- (E) a dedução de custos de investimento individuais em relação aos possíveis rendimentos.

## 15

O termo "each" (segundo parágrafo) refere-se a

- (A) "Nash equilibrium".
- (B) "unexpected results".
- (C) "common good".
- (D) "N interested parties".
- (E) "share".

## 16

Segundo o estudo, a "Tragédia dos Comuns" pode trazer mais prejuízos do que aqueles considerados inicialmente e pode levar aos seguintes efeitos:

- (A) Exploração excessiva do recurso e danos aos investidores.
- (B) Acesso descontrolado aos bens e ausência de lucratividade.
- (C) Pobreza catastrófica e não cooperação entre participantes.
- (D) Evasão do mercado e competição desregulada.
- (E) Redução dos investimentos e corte dos retornos financeiros.

### TEXTO PARA AS QUESTÕES 17 A 20

#### Quantum breakthrough could revolutionise computing

Computer scientists have been trying to make an effective quantum computer for more than 20 years. Firms such as Google, IBM and Microsoft have developed simple machines. But, according to Prof. Winfried Hensinger, who led the research at Sussex University, the new development *paves the way* for systems that can solve complex real world problems that the best computers we have today are incapable of.

"Right now we have quantum computers with very simple microchips," he said. "What we have achieved here is the ability to realise extremely powerful quantum computers capable of solving some of the most important problems for industries and society."

Currently, computers solve problems in a simple linear way, one calculation at a time. In the quantum realm, particles can be in two places at the same time and researchers want to harness this property to develop computers that can do multiple calculations all at the same time.

Quantum particles can also be millions of miles apart and be strangely connected, mirroring each other's actions instantaneously. Again, that *could* also be used to develop much more powerful computers.

One stumbling block has been the need to transfer quantum information between chips quickly and reliably: the information degrades, and errors are introduced.

But Prof. Hensinger's team has made a breakthrough, published in the journal Nature Communications, which may have overcome that obstacle.

The team developed a system able to transport information from one chip to another with a reliability of 99.999993% at record speeds. That, say the researchers, shows that in principle chips could be slotted together to make a more powerful quantum computer.

GHOSH, Pallab. Quantum breakthrough could revolutionise computing. BBC News (online). 08 Feb. 2023 (adaptado)..



**17**

A expressão idiomática “paves the way” (primeiro parágrafo) pode ser traduzida como

- (A) dá a forma.
- (B) denota o estilo.
- (C) especifica o jeito.
- (D) fortalece a maneira.
- (E) abre o caminho.

**18**

O verbo modal “could” (quarto parágrafo) tem efeito de

- (A) capacidade.
- (B) possibilidade.
- (C) permissão.
- (D) solicitação.
- (E) necessidade.

**19**

De acordo com o texto, a vantagem dos computadores quânticos é

- (A) utilizar microchips para realizar cálculos diversos.
- (B) produzir contagens industriais de forma mais veloz.
- (C) desenvolver uma operação linear por vez.
- (D) possuir partículas em dois locais simultaneamente.
- (E) apresentar conexões e interações entre si.

**20**

Segundo o texto, a equipe do Prof. Hensinger desenvolveu um grande avanço nas pesquisas, pois possibilitou

- (A) transferir informação quântica entre chips com rapidez e segurança.
- (B) identificar onde a informação se degrada e onde surgem erros.
- (C) acelerar a comunicação sobre falhas entre computadores.
- (D) gerar um sistema de compartilhamento de dados em velocidade recorde.
- (E) armazenar computadores quânticos mais poderosos.

**21**

Uma loja especializada em bicicletas anuncia um desconto de 25% em um determinado modelo. O valor com desconto de tal modelo é R\$ 1.980,00. O preço original de tal bicicleta é

- (A) R\$ 2.200,00
- (B) R\$ 2.450,00
- (C) R\$ 2.640,00
- (D) R\$ 2.830,00
- (E) R\$ 2.970,00

**22**

Em uma progressão aritmética, de razão não nula e termos estritamente positivos, os 1º, 10º e 20º termos formam, nesta ordem, uma progressão geométrica. A razão desta PG é

- (A)  $\frac{7}{8}$
- (B)  $\frac{9}{10}$
- (C)  $\frac{10}{9}$
- (D)  $\frac{12}{10}$
- (E)  $\frac{13}{11}$

**23**

Considere  $A = \cos x + \cos y$  e  $B = \sin x - \sin y$ . Se  $x + y = 135^\circ$ , o valor  $3A^2 + 3B^2$  é

- (A) 1
- (B)  $6 - 3\sqrt{2}$
- (C) 3
- (D)  $3\sqrt{2}$
- (E) 6

**24**

Em um estabelecimento comercial trabalham 5 entregadores de mercadorias. Uma das maneiras de realizar a entrega é utilizando a única bicicleta que pertence ao estabelecimento. Os entregadores são: Alberto (20 anos), Raul (22 anos), Jair (25 anos), Luís (28 anos) e Gerson (29 anos). Cada um pode utilizar a única bicicleta em um só dia da semana. Sabe-se que:

- Raul a usa um dia depois de Luís.
- Jair a usa às sextas-feiras.
- Nem Luís e nem Raul podem usá-la às quartas-feiras.
- Alberto a usa na quinta-feira.

A soma das idades dos entregadores que a usam às segundas e quartas-feiras é:

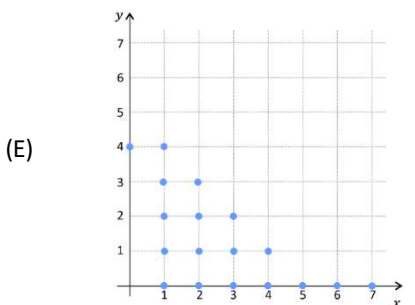
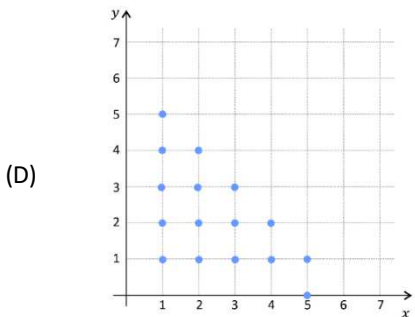
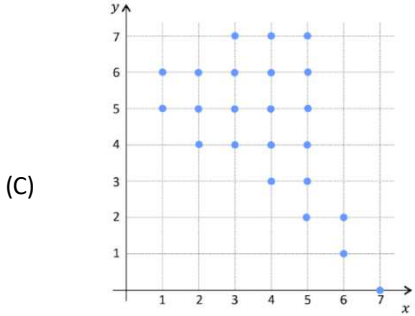
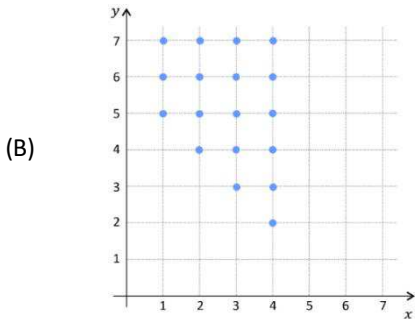
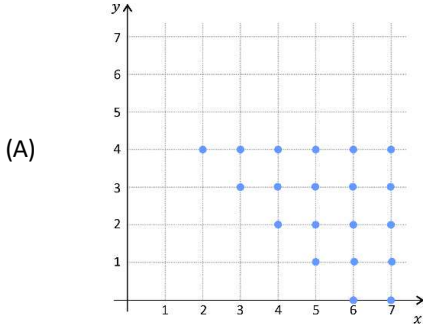
- (A) 49
- (B) 51
- (C) 53
- (D) 54
- (E) 57





**25**

Retiram-se ao azar pelo menos 6 bolas de uma caixa que possui no máximo 7 bolas azuis e no máximo 4 bolas amarelas. Se os pares possíveis forem representados no plano cartesiano, em que o *eixo x* corresponde às bolas azuis e o *eixo y* às bolas amarelas, o gráfico que melhor representa os possíveis pares que podem ser retirados é:



**26**

Sobre duas retas paralelas  $r$  e  $s$  são marcados 13 pontos, sendo a maioria deles na reta  $r$ . Com estes 13 pontos como vértices, constroem-se todos os triângulos e todos os quadriláteros convexos possíveis. Sabendo-se que a razão entre o número de quadriláteros e o número de triângulos é  $\frac{14}{11}$ , o dobro do número dos pontos que pertencem a  $r$  somado ao triplo dos que pertencem a  $s$  será

- (A) 28
- (B) 29
- (C) 30
- (D) 31
- (E) 32

**27**

Uma concessionária de automóveis planeja contratar três empresas de pesquisa de opinião para realizar uma pesquisa com 950 clientes por WhatsApp, 600 clientes por e-mail e 400 pessoalmente. Tem-se que, por hora:

- A empresa A pode realizar 30 consultas por WhatsApp, 10 por e-mail e 5 pessoalmente.
- A empresa B pode realizar 10 consultas por WhatsApp, 20 por e-mail e 10 pessoalmente.
- A empresa C pode realizar 20 consultas por WhatsApp, 10 por e-mail e 15 pessoalmente.

Se  $a, b$  e  $c$  indicarem os números de horas contratadas das empresas A, B e C, respectivamente, para se obter o número exato de consultas requeridas, então

- (A)  $a = 30h, b = 20h$  e  $c = 15h$
- (B)  $a = 25h, b = 15h$  e  $c = 10h$
- (C)  $a = 20h, b = 20h$  e  $c = 10h$
- (D)  $a = 20h, b = 15h$  e  $c = 10h$
- (E)  $a = 25h, b = 20h$  e  $c = 20h$

**28**

Se o polinômio  $3x^2 - 9x + 7$  for decomposto em uma diferença de dois polinômios cúbicos na forma  $(x - a)^3 - (x - b)^3$ , o valor de  $a + b$  será

- (A) 1
- (B) 3
- (C) 4
- (D) 5
- (E) 6



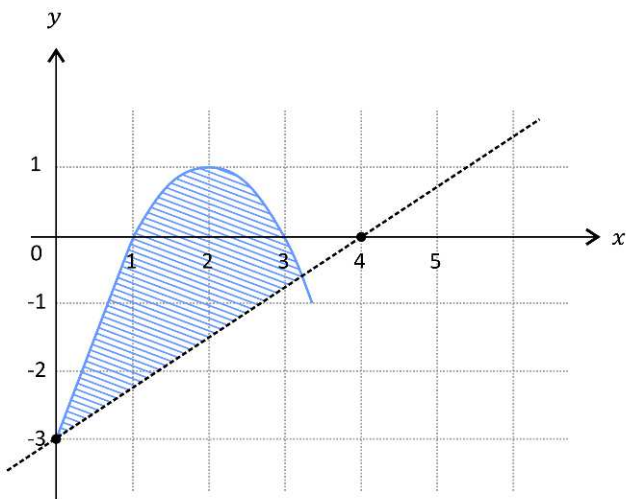
**29**

Considere a matriz  $A = \begin{pmatrix} -2 & 4 \\ 6 & -8 \end{pmatrix}$ . Sabendo-se que a sua inversa é  $A^{-1} = \begin{pmatrix} 1 & \frac{1}{2} \\ m & n \end{pmatrix}$ , o valor de  $m + n$  será

- (A) 1
- (B)  $\frac{1}{2}$
- (C) -1
- (D)  $-\frac{1}{2}$
- (E) 0

**30**

Observe a figura apresentada.



O sistema de inequações que descreve a região hachurada é

- (A)  $x^2 - 4x + 3 \geq y$  e  $4y - 3x < -12$
- (B)  $-x^2 + 5x - 4 \leq y$  e  $4y - 3x > -12$
- (C)  $-x^2 + 4x - 3 \geq y$  e  $-4y + 3x < 12$
- (D)  $x^2 - 3x + 2 \geq y$  e  $4y - 3x < -12$
- (E)  $x^2 + 4x + 3 \geq y$  e  $4y - 3x > -12$

**31**

O valor do limite  $\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2}$  é

- (A) -1
- (B) 0
- (C) 1
- (D)  $+\infty$
- (E)  $\frac{1}{2}$

**32**

O valor do limite

$$\lim_{x \rightarrow +\infty} \frac{(x-1)(x-2)(x-3)(x-4)(x-5)}{(5x-1)^5}$$

é

- (A) 0
- (B)  $5^{-5}$
- (C)  $5^{-3}$
- (D)  $5^{-1}$
- (E)  $+\infty$

**33**

Seja  $f: \mathbb{R} \rightarrow \mathbb{R}$  a função dada por  $f(x) = e^{(x-2)} - 3x$ , pode-se afirmar que  $f$

- (A) não intersecta o eixo das abscissas
- (B) intersecta o eixo das abscissas no intervalo  $[-5, -2]$
- (C) intersecta o eixo das abscissas no intervalo  $[-3, -2]$
- (D) intersecta o eixo das abscissas nos intervalos  $[0, 2]$  e  $[3, 5]$
- (E) intersecta o eixo das abscissas nos intervalos  $[6, 8]$  e  $[9, 10]$

**34**

Se  $f: \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}$  for a função definida por

$$f(x) = \frac{x}{x-1}$$

então,

- (A)  $\lim_{x \rightarrow +\infty} f(x) = +\infty$
- (B)  $f$  não tem ponto de inflexão
- (C)  $f$  é uma função crescente no intervalo  $]1, +\infty[$
- (D)  $f$  tem concavidade para cima no intervalo  $]-\infty, 1[$
- (E)  $f$  tem um ponto de mínimo no intervalo  $]1, 2[$



**35**

Seja  $f: \mathbb{R} \rightarrow \mathbb{R}$  a função definida por

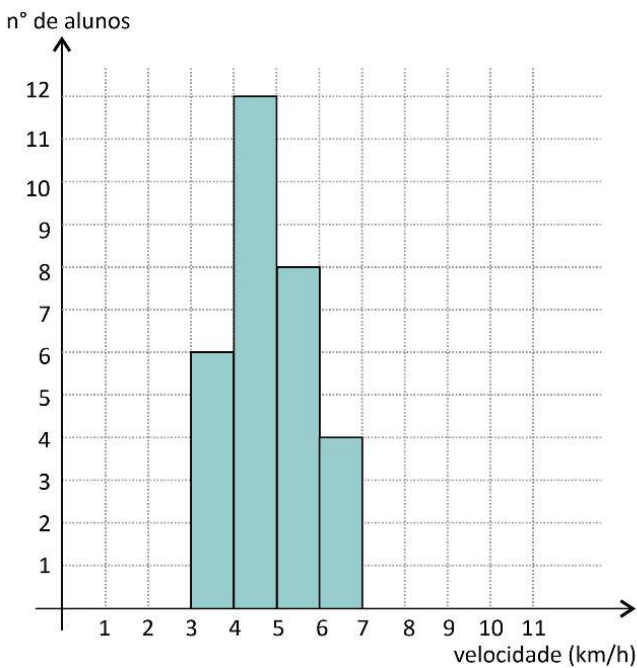
$$f(x) = e^{(-x^3+3x)}$$

Pode-se afirmar que

- (A)  $x = 3$  é um ponto de mínimo da  $f$
- (B)  $f$  é uma função crescente no intervalo  $]1, +\infty[$
- (C)  $\lim_{x \rightarrow -\infty} f(x) = 0$
- (D)  $f$  é uma função crescente no intervalo  $] -1, 1[$
- (E)  $\lim_{x \rightarrow -\infty} f(x) = +\infty$

**36**

Um professor de academia, em um treino de caminhada ao ar livre, registra a velocidade de caminhada, em km/h, de cada aluno ao longo de uma manhã. Ele elabora o histograma apresentado.



A mediana das velocidades de caminhada dos alunos no histograma em km/h é

- (A) 3,95
- (B) 4,15
- (C) 4,35
- (D) 4,55
- (E) 4,75

**37**

O valor da integral  $\int_{-1}^1 \frac{x dx}{\sqrt{5-4x}}$  é

- (A)  $\frac{1}{6}$
- (B)  $\frac{1}{3}$
- (C)  $\frac{1}{2}$
- (D) 1
- (E)  $\frac{3}{2}$

**38**

Seja  $f: \mathbb{R} \rightarrow \mathbb{R}$  a função dada por

$$f(x) = \begin{cases} (x-p)^2 - 3, & \text{se } x \geq 0 \\ q \sin 2x, & \text{se } x < 0 \end{cases}$$

Os possíveis valores que a expressão  $-p + q$  assume quando  $f$  for uma função contínua e diferenciável em  $x = 0$  são

- (A)  $-\sqrt{3}$  ou  $\sqrt{3}$
- (B)  $-1$  ou  $1$
- (C)  $-2\sqrt{3}$  ou  $2\sqrt{3}$
- (D)  $-3\sqrt{3}$  ou  $3\sqrt{3}$
- (E)  $-4\sqrt{3}$  ou  $4\sqrt{3}$

**39**

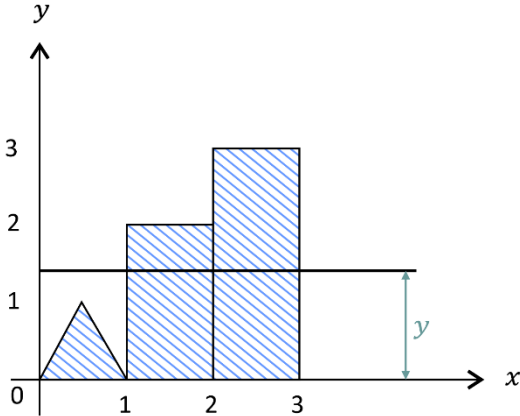
Considere  $\beta$  um valor real maior do que zero. O valor da área limitada pelas curvas  $y^2 = \beta x$ ;  $x + 3y = 4\beta$  e pelas retas  $x = 0$  e  $x = \beta$  é

- (A)  $\frac{1}{2} \beta^2$
- (B)  $\frac{1}{3} \beta^2$
- (C)  $\frac{1}{4} \beta^2$
- (D)  $\frac{1}{6} \beta^2$
- (E)  $\frac{1}{12} \beta^2$



40

A figura apresentada é formada por um triângulo de base 1 e altura 1 e dois retângulos de base 1 e alturas 2 e 3 respectivamente. A função  $G(y)$  representa a área da parte da figura que está compreendida entre o eixo  $x$  e a reta paralela a tal eixo passando por  $(0, y)$ .



As expressões analíticas de  $G$ , em função de  $y$ , em cada intervalo dado são

(A)  $G(y) = \begin{cases} 3y - \frac{y^2}{2}, & 0 \leq y \leq 1 \\ \frac{1}{2} + 2y, & 1 < y \leq 2 \end{cases}$

(B)  $G(y) = \begin{cases} \frac{1}{2} + 2y, & 1 < y \leq 2 \\ \frac{9}{2} + y, & 2 < y \leq 3 \end{cases}$

(C)  $G(y) = \begin{cases} 3y - \frac{y^2}{2}, & 0 \leq y \leq 1 \\ -\frac{1}{2} + y, & 1 < y \leq 2 \end{cases}$

(D)  $G(y) = \begin{cases} -\frac{1}{2} + y, & 1 < y \leq 2 \\ -\frac{5}{2} + 2y, & 2 < y \leq 3 \end{cases}$

(E)  $G(y) = \begin{cases} -\frac{5}{2} + 2y, & 2 < y \leq 3 \\ \frac{11}{2} + y, & 3 < y < +\infty \end{cases}$

41

Ao se considerar a modelagem de dados em um ambiente de *data warehouse*, qual das seguintes afirmações descreve corretamente vantagens e/ou desvantagens do esquema “estrela” (*star schema*) e do esquema “flocos de neve” (*snowflake schema*)?

- (A) O esquema “flocos de neve”, devido à sua estrutura mais simples, pode exigir menos esforço para manutenção e modificações, diminuindo a complexidade do desenvolvimento de consultas e relatórios.
- (B) Diferentemente do esquema “flocos de neve”, o esquema “estrela” reduz o espaço de armazenamento dos dados ao unificar tabelas dimensionais relacionadas, pois evita duplicação desnecessária de dados.
- (C) O esquema “flocos de neve” é mais flexível e escalável, embora requeira mais junções entre tabelas para consultar os dados, o que pode aumentar a complexidade e afetar negativamente o desempenho de consulta.
- (D) No esquema “estrela”, as tabelas de dimensões não são normalizadas para poder reduzir a redundância e melhorar a consistência dos dados, enquanto o esquema flocos de neve mantém uma estrutura mais normalizada.
- (E) A principal diferença entre o esquema “estrela” e o esquema “flocos de neve” está na modelagem das dimensões, pois, no esquema estrela, ocorrem ligações entre as tabelas dimensionais, e, no esquema “flocos de neve”, não.

42

Um modelo de regressão linear simples foi ajustado sobre um conjunto de pares  $(x_1, y_1) \dots (x_5, y_5)$ , com o objetivo de se prever o valor da variável  $y$ , dado o valor de  $x$ .

Denotamos por  $\hat{y}_i$  o valor predito de  $y_i$ , dado o valor de  $x_i$ , e denotamos por  $\bar{y}$  a média dos valores de  $y_1, \dots, y_5$ .

Suponha que, após o ajuste do modelo, os seguintes indicadores de ajuste tenham sido obtidos:

$$TSS = \sum_{i=1}^5 (y_i - \bar{y})^2 = 16.$$

$$RSS = \sum_{i=1}^5 (y_i - \hat{y}_i)^2 = 4;$$

Com base nestas informações, qual é o valor do coeficiente de determinação,  $R^2$ ?

- (A) 0,20.
- (B) 0,25.
- (C) 0,50.
- (D) 0,75.
- (E) 0,80.



43

Sobre a krigagem, no contexto de análise de dados georreferenciados, a seguinte afirmação é correta:

- (A) É uma técnica concebida para a predição de valores de uma ou mais variáveis em pontos quaisquer de uma área geográfica, a partir de uma amostra de pontos previamente observados nessa área; tal predição leva em consideração a estrutura de covariância espacial dos pontos observados.
- (B) É uma técnica concebida para a suavização dos valores de uma ou mais variáveis em um conjunto de dados georreferenciados; baseada no LOWESS (*locally weighted estimated scatterplot smoothing*), a função de atenuação leva em consideração a estrutura de covariância espacial dos pontos observados.
- (C) É uma técnica de conversão entre sistemas de referência de coordenadas, concebida para permitir a incorporação de variáveis numéricas em um Sistema de Informações Geográficas - o que não é possível, por exemplo, com o sistema geodésico mundial WGS84.
- (D) É uma técnica baseada no *Fuzzy c-means*, concebida para detecção de agrupamentos em dados georreferenciados; o cálculo das distâncias entre os pontos observados e os centroides dos agrupamentos combina o vetor original de cada ponto com sua coordenada geográfica, explorando, dessa forma, a estrutura de covariância espacial dos dados.
- (E) É uma técnica baseada no LOWESS (*locally weighted estimated scatterplot smoothing*), concebida para a detecção e remoção de *outliers* em dados georreferenciados; o limiar para determinar se um valor é ou não um *outlier* é calculado levando-se em consideração a estrutura de covariância espacial dos dados.

44

Considerando a área de Processamento de Linguagem Natural, assinale a alternativa correta.

- (A) LDA (*Latent Dirichlet Allocation*) é uma técnica utilizada em modelagem de tópicos que considera que cada texto é predominantemente de um único tópico a ser predito por aprendizado supervisionado.
- (B) Representações de texto baseadas em n-gramas possuem a desvantagem de criarem um espaço de características mais esparsas que a estratégia de *bag of words* simples (uma só palavra).
- (C) BERT (*Bidirecional Encoder Representations from Transformers*) é um exemplo de classificador não-supervisionado baseado em propriedades sonoras independentes entre si para classificação de sentimentos a partir de textos.
- (D) *Word embeddings* é um tipo de representação vetorial em que o texto todo é representado como um vetor n-dimensional, sendo n o número total de palavras do dicionário linguístico adotado.
- (E) Tf-idf é um modelo de transformadores (do inglês *transformers*) baseado em redes neurais profundas generativas.

45

O esquema parcial de um Banco de Dados Relacional de uma Universidade é apresentado a seguir, composto pelas relações Estudante, Disciplina, TurmaDisciplina e HistoricoEscolar. Os nomes entre parênteses definem os atributos armazenados para cada relação; atributos sublinhados indicam chave primária e atributos em negrito indicam chave estrangeira. Atributos que compõem chaves estrangeiras são identificados com o mesmo nome dos atributos que compõem chaves primárias às quais fazem referência. Considere a execução do seguinte comando em *Structured Query Language* (SQL) padrão:

Estudante (e-codigo, e-nome, e-curso)

Disciplina (d-codigo, d-nome, d-quantidade-aulas)

TurmaDisciplina (td-codigo, **d-codigo**, td-ano, td-semester, td-local-aula)

HistoricoEscolar (**e-codigo**, **td-codigo**, he-nota-final, he-frequencia)

**Comando SQL executado:**

```
select
e.e-nome, d.d-nome, he.he-nota-final
from
Estudante e, Disciplina d, TurmaDisciplina td,
HistoricoEscolar he
where
he.e-codigo = e.e-codigo and
he.td-codigo = td.td-codigo and
td.d-codigo = d.d-codigo and
td.td-ano = 2023 and
td.td-semester = 1 and
(he.td-codigo, he.he-nota-final) in
(select he2.td-codigo, max(he2.he-nota-final)
from HistoricoEscolar he2
group by he2.td-codigo);
```

É correto afirmar que o comando mostra o nome do aluno, o nome da disciplina e a nota final

- (A) de um aluno que obteve a maior nota final de cada uma das turmas oferecidas no primeiro semestre de 2023.
- (B) de todos os alunos que obtiveram a maior nota final de cada uma das turmas oferecidas no primeiro semestre de 2023.
- (C) de todos os alunos que obtiveram a maior nota final de todas as turmas oferecidas no ano de 2023.
- (D) do aluno que obteve a maior nota final entre todas as turmas oferecidas no primeiro semestre de 2023.
- (E) somente do primeiro aluno gravado no banco de dados que obteve a maior nota final em cada uma das turmas oferecidas no primeiro semestre de 2023.



**46**

A visualização de dados vem se tornando um mecanismo indispensável em organizações para tomada de decisão em diversos níveis hierárquicos gerenciais. Ferramentas de visualização de dados, também denominadas ferramentas analíticas, tais como *Power BI* e *Tableau*, têm sido amplamente divulgadas e empregadas para apoiar a decisão baseada em dados. Essas ferramentas possibilitam a importação

- (A) de dados a partir de bancos de dados e a criação de painéis de visualização somente estáticos a partir desta importação.
- (B) somente a partir de bancos de dados relacionais e a criação de painéis de visualização somente estáticos a partir desta importação.
- (C) de dados a partir de bancos de dados e a criação de painéis de visualização dinâmicos a partir desta importação.
- (D) de dados somente a partir de bancos de dados relacionais e a criação de painéis de visualização dinâmicos a partir desta importação.
- (E) somente de dados armazenados por ferramentas do mesmo fabricante e a criação de painéis de visualização dinâmicos a partir desta importação.

**47**

Sobre técnicas de agrupamento (clusterização), assinale a alternativa correta.

- (A) Algoritmos de agrupamento hierárquico somente podem ser utilizados quando se sabe o número de classes presentes nos dados.
- (B) Se os dados pertencerem a mais de duas classes, os agrupamentos são identificados utilizando um algoritmo supervisionado de classificação multiclasse.
- (C) K-médias é um exemplo de algoritmo de agrupamento particional no qual o número de classes que se pretende encontrar deve ser informado.
- (D) PCA (análise de componentes principais) é um exemplo de algoritmo de agrupamento hierárquico no qual as classes principais são identificadas primeiro.
- (E) K-vizinhos-mais-próximos é uma medida de distância entre grupos utilizada em algoritmos de agrupamentos particionais que tendem a encontrar grupos compactos.

**48**

Um conceito importante em ciência de dados é o de *Big Data*. As três características mais importantes deste tipo de dados são conhecidas como os três V's, que são:

- (A) Velocidade, vitalidade, volume.
- (B) Volume, visibilidade, validade.
- (C) Variedade, virtual, vitalidade.
- (D) Verossímil, viável, visibilidade.
- (E) Variedade, volume, velocidade.

**49**

Sobre redução de dimensionalidade e seleção de características, assinale a alternativa correta.

- (A) PCA (análise de componentes principais) é uma técnica supervisionada de redução de dimensionalidade.
- (B) O algoritmo de indução de uma árvore de decisão já realiza seleção de características durante o treinamento.
- (C) Todas as técnicas de redução de dimensionalidade sempre selecionam um subconjunto das características originais.
- (D) *Oversampling* é uma técnica de seleção de características baseada na priorização das características mais frequentes.
- (E) Não é possível realizar redução de dimensionalidade sem uma amostra de treinamento.

**50**

Sobre o algoritmo SVM (máquinas de vetores de suporte), pode-se afirmar que são classificadores

- (A) não-lineares multiclasse que usam funções kernel para mapear um espaço de características separado para cada classe presente na amostra de treinamento.
- (B) de texto que usam a representação vetorial de palavras e que usam diferentes funções kernel para identificar o tópico central do texto.
- (C) que não necessitam de uma etapa de treinamento, pois usam funções kernel para identificar a classe mais próxima da entrada.
- (D) lineares binários que usam funções kernel para manipular os dados em um espaço que pode ou não ser diferente do espaço de características original.
- (E) não supervisionados para criação de agrupamentos que usam funções kernel para identificar k centroides, um de cada classe, sendo k informado inicialmente.

**51**

Considere que você possua um *dataset* contendo 100 instâncias de uma classe A e 120 instâncias de uma classe B. Você utilizou 80% das instâncias de cada classe deste *dataset* para treinar um classificador, e o utilizou para prever a classe de todas as 220 instâncias do *dataset*. Curiosamente, seu classificador acertou a classe de todas as instâncias que foram utilizadas no treinamento do classificador, mas acertou apenas cerca de 50% das instâncias que não foram usadas no treinamento. Este é um cenário típico que indica que você deveria ter aplicado qual técnica no seu processo de treinamento?

- (A) Normalização.
- (B) Balanceamento de classes.
- (C) Remoção de *outliers*.
- (D) Regularização.
- (E) Clusterização.



52

Considere um classificador binário treinado para classificar dados de um e-mail em spam ou não-spam. A seguinte matriz de confusão foi obtida após a aplicação desse classificador binário em uma amostra de teste:

|            |          | Valor predito |          |
|------------|----------|---------------|----------|
|            |          | Spam          | Não-spam |
| Valor Real | Spam     | 40            | 20       |
|            | Não-spam | 10            | 130      |

Utilizando a matriz acima, quais os valores de acurácia, precisão e revocação (recall ou sensibilidade), respectivamente?

- (A) 0,40; 1,30; 0,30.
- (B) 0,85; 0,93; 0,20.
- (C) 0,15; 0,80; 0,30.
- (D) 0,40; 0,93; 0,67.
- (E) 0,85; 0,80; 0,67.

53

Suponha que você possua dados de uma grande rede de lojas e que precise criar um classificador para prever se um determinado cliente comprará ou não um determinado produto. Para isso, a empresa disponibilizou uma amostra contendo dados de 10.000 clientes que compraram o produto e de 10.000 clientes que não compraram o produto. Dentre os vários atributos (ou características) de cada cliente estão: faixa de renda familiar (um valor discreto de 1 a 5), idade, o valor total de produtos já comprados da empresa (somando todas as compras já realizadas), e se comprou ou não o produto (0 = não, 1 = sim) durante uma campanha. Antes de executar um algoritmo de treinamento do classificador, um procedimento importante de pré-processamento a ser realizado sobre esses dados é:

- (A) Balanceamento de classes.
- (B) Normalização das características.
- (C) Calibração de hiperparâmetros.
- (D) *Overfitting*.
- (E) Lematização.

54

Em Mineração de Dados, após o pré-processamento dos dados, inicia-se a fase de análises. Para esta fase há análises descritivas e análises preditivas. Quais das tarefas a seguir são todas análises descritivas?

- (A) Identificação de *outliers*, regressão e clusterização.
- (B) Classificação supervisionada, identificação de regras de associação e análise de correlação.
- (C) *Underfitting*, clusterização e sumarização.
- (D) Visualização, sumarização e identificação de regras de associação.
- (E) Sumarização, regressão e visualização.

55

Sobre redes neurais convolucionais profundas (CNNs), assinale a alternativa correta.

- (A) A aplicação de *dropout* é importante para ajustar o formato dos dados para serem processados pela camada totalmente conectada.
- (B) São tradicionalmente adequadas para processar dados textuais devido à sua arquitetura baseada em *word embeddings* dinâmicos.
- (C) *Data augmentation* é o processo de transferência de aprendizado em que o *fine-tuning* é realizado nas primeiras camadas convolucionais da rede.
- (D) As camadas de convolução realizam extração de características enquanto as camadas de *pooling* reduzem a dimensionalidade.
- (E) CNNs podem ser usadas como extratores de características para classificadores SVM porque compartilham os mesmos tipos de kernels.

56

Suponha que você possua um *dataset* que represente uma amostra de treinamento balanceada de 500.000 instâncias para classificação binária. Devido à demanda de tempo de processamento dos algoritmos de indução de classificadores que você usará, você definiu que cada treinamento deve ser realizado apenas por 40.000 instâncias e cada teste sobre 10.000 instâncias. Uma vez que 1) cada treinamento e teste demoram juntos 2 horas, 2) você tem disponível apenas 20 horas para executar a estimação de desempenho do algoritmo usado em seus dados, e 3) você quer fazer o máximo de treinamentos/testes possível nessas 20 horas, uma boa estratégia a ser aplicada sobre seu *dataset* para a estimação de desempenho é:

- (A) validação cruzada 10-vezes (ou 10X) estratificada a partir de todo o *dataset* original.
- (B) *leave-one-out* de forma a manter o balanceamento entre as classes em cada uma das 10 execuções.
- (C) 10 *hold-outs* balanceados independentes, cada um contendo 50.000 instâncias sorteadas do *dataset*.
- (D) *grid-search* balanceado com dimensão total = 40.000 x 10.000 x 10 com *bootstrap*.
- (E) 10 etapas de resubstituição do *dataset* original já balanceado e sem *bootstrap*.

57

Dados abertos são dados que

- (A) podem ser livremente usados, reusados e redistribuídos.
- (B) podem ser usados, reusados, mas não redistribuídos.
- (C) podem ser gratuitamente obtidos, mas não podem ser redistribuídos sem o pagamento de uma taxa.
- (D) podem ser livremente usados, reusados e redistribuídos somente se não estiverem protegidos por senha.
- (E) podem ser gratuitamente obtidos apenas os campos não estruturados.



**58**

Considere o seguinte código na linguagem de programação Python:

```
def misterio(a):  
    aux = 0.0  
    for v in a:  
        if v < aux:  
            aux = v  
    a.append(aux)  
    c = (aux, a)  
    return c  
  
b = [2, 4, 1, -5, 10]  
resp = misterio(b)  
print (resp[1])
```

Qual é a saída que será impressa após a execução deste programa?

- (A) [2, 4, 1, -5, 10]
- (B) 4
- (C) [2, 4, 1, -5, 10, -5]
- (D) -5
- (E) Erro de compilação

---

**59**

Sobre anomalias (*outliers*), assinale a alternativa correta.

- (A) As anomalias normalmente residem em regiões densas do espaço de características.
- (B) Dependendo do problema a ser tratado, as anomalias devem ser eliminadas ou são justamente o que se pretende modelar.
- (C) Anomalias nos dados de treinamento não interferem no desempenho de um classificador supervisionado.
- (D) Algoritmos de tokenização frequentista podem ser utilizados para a detecção de anomalias.
- (E) A aplicação de uma normalização do tipo min-max é suficiente para eliminar anomalias.

---

**60**

Suponha que você precisa analisar *posts* em uma rede social sobre opiniões de clientes da sua empresa. O conjunto de textos desses *posts* são considerados dados

- (A) categóricos.
- (B) estruturados.
- (C) intervalares.
- (D) nominais.
- (E) não estruturados.





## QUESTÕES DISSERTATIVAS

### QUESTÃO 01:

Validação cruzada *k-vezes*, *leave-one-out* e *bootstrap* são técnicas aplicadas em aprendizado de máquina e mineração de dados. Explique essas três técnicas, incluindo: a) a que objetivo elas se destinam; b) a descrição do funcionamento de cada uma delas; c) as vantagens e as desvantagens de cada técnica.

### QUESTÃO 02:

Dentro do contexto de aprendizado de máquina, explique a redução de dimensionalidade e a seleção de características incluindo os seguintes tópicos: a) o que são os processos de redução de dimensionalidade e seleção de características; b) o porquê destes processos serem importantes em um *pipeline* de aprendizado de modelos (cite três motivos).

### Instruções:

- Na resolução da questão, não ultrapasse o quadro destinado à resposta.
- Em hipótese alguma o candidato deverá se identificar no campo destinado ao preenchimento da questão dissertativa. Receberão nota zero os textos que permitirem, por qualquer modo, a identificação do candidato.



v.2

**Concursos DRH 2023**  
1ª Fase – Objetiva e Dissertativa

0/0

1

1/100

