

PROJETO

INTEGRADO



PROJETO INTEGRADO

BOAS-VINDAS

Prezado(a) aluno(a),

Seja bem-vindo(a) ao Projeto Integrado! Esta é uma oportunidade de aplicar os conhecimentos adquiridos nas disciplinas do semestre em uma situação realista e desafiadora que terá como foco Análise, Otimização e Segurança de Processos de Atendimento ao Cliente com Aplicação de NLP, Probabilidade Estatística e Pesquisa Operacional. Esperamos que você se envolva ativamente, desenvolvendo competências essenciais para sua formação profissional.

Você terá a oportunidade de conectar teoria e prática, voltando-se para os desafios do mercado de trabalho.

Desejamos a você muito sucesso no desenvolvimento dessa atividade!

OBJETIVOS DE APRENDIZAGEM

Você já deve ter percebido que, ao longo do curso, cada disciplina apresenta conceitos, ferramentas e práticas que, quando combinadas, formam o conhecimento que define o seu perfil profissional. O Projeto Integrado foi criado exatamente para unir tudo isso: é o momento em que você aplica, de forma prática e reflexiva, aquilo que aprendeu durante o semestre, demonstrando como a teoria se transforma em solução para desafios reais da sua área de formação.

Essa atividade é um componente curricular obrigatório e faz parte do seu processo de aprendizagem. Caso você não realize o projeto ou não atinja a pontuação necessária, será reprovado. Por isso, é fundamental dedicar-se com responsabilidade, planejamento e empenho. Afinal, o Projeto Integrado é uma oportunidade única de consolidar os conhecimentos adquiridos e mostrar o quanto você evoluiu academicamente. A pontuação mínima para aprovação deve ser verificada no Manual da Avaliação Continuada.

Mais do que uma avaliação, o Projeto Integrado é uma experiência de integração e construção de saberes. Ele propõe que você desenvolva um olhar crítico e criativo diante de uma situação-problema, analisando o contexto, aplicando conceitos das disciplinas do semestre e apresentando uma proposta de solução prática e bem fundamentada. Trata-se de um espaço para exercitar a autonomia intelectual, o pensamento científico e a capacidade de resolução de problemas — habilidades essenciais no mundo do trabalho e valorizadas em qualquer carreira.

Durante o desenvolvimento do projeto, você será convidado a realizar um estudo dirigido, com orientação a distância e acompanhamento dos professores, que o guiarão na aplicação dos conteúdos de

PROJETO INTEGRADO

forma interdisciplinar. Isso significa que as disciplinas do semestre se complementam e dialogam entre si, contribuindo para que o trabalho final seja coeso, relevante e alinhado à realidade da profissão.

O Projeto Integrado também tem como propósito aproximar você do saber científico, estimulando a produção de conhecimento acadêmico com base em normas da ABNT, algo indispensável à formação de um profissional ético, competente e preparado para lidar com dados, evidências e argumentos sólidos. Ao final, o mais importante é compreender que o Projeto Integrado não é apenas uma tarefa — é um processo de aprendizagem que une conhecimento, criatividade e prática profissional. Ele representa o seu compromisso com a formação de qualidade e o seu protagonismo como estudante.

Cada etapa concluída é uma oportunidade de aprender mais sobre si, sobre a sua área de atuação e sobre como transformar conhecimento em ação.

ORIENTAÇÕES INICIAIS

A partir da análise do contexto inicial apresentado, você, como futuro profissional, deverá desenvolver uma atividade fundamentada nas disciplinas estudadas ao longo deste semestre.

Essa atividade visa aprimorar suas habilidades de leitura, interpretação e aplicação dos conhecimentos adquiridos, proporcionando uma compreensão mais profunda e crítica dos conteúdos abordados.

Uma empresa nacional de varejo digital, recebe diariamente milhares de mensagens de clientes pelos canais de atendimento: chat, e-mail e WhatsApp.

Com o aumento da demanda, surgiram problemas:

- Altos **tempos de espera**
- Dificuldade em **classificar automaticamente** os tipos de mensagens
- Falta de **análise estatística** dos volumes de atendimento
- Impossibilidade de prever **picos de demanda**
- Ausência de **simulações operacionais** para otimizar a fila
- Riscos relacionados ao **tratamento de dados pessoais** nas mensagens

A diretoria decidiu iniciar um **projeto piloto de Ciência de Dados** para:

- Classificar automaticamente mensagens com **NLP**.
- Analisar padrões estatísticos dos atendimentos.
- Simular a fila de atendimento como um sistema de **eventos discretos**.
- Avaliar riscos de segurança e cumprir LGPD.

PROJETO INTEGRADO

Sua equipe precisa desenvolver uma **POC (Prova de Conceito)** que demonstre a viabilidade do projeto.

ATIVIDADE A SER REALIZADA

A partir da análise do contexto inicial apresentado, você, como futuro profissional, deverá desenvolver uma atividade fundamentada nas disciplinas estudadas ao longo deste semestre.

ATIVIDADE PRÁTICA NO GOOGLE COLAB

Copie e cole o código **completo** abaixo no Colab.

Você deve executar **cada etapa**, interpretar resultados e responder às questões solicitadas.

```
import pandas as pd
import numpy as np
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
import matplotlib.pyplot as plt
import random

np.random.seed(42)

# =====
# 1. GERAÇÃO DE DADOS FICTÍCIOS DE ATENDIMENTO
# =====

tipos = ["reclamação", "elogio", "dúvida", "cancelamento"]
mensagens_exemplo = {
    "reclamação": ["Produto chegou quebrado", "Não recebi o pedido",
"atendimento ruim"],
    "elogio": ["Gostei muito do produto", "Entrega rápida", "Atendente muito
educado"],
    "dúvida": ["Qual o prazo de entrega?", "Como faço para trocar?", "Tem
garantia?"],
    "cancelamento": ["Quero cancelar o pedido", "Desejo reembolso", "Cancelar
assinatura"]
}

dados = []
```

PROJETO INTEGRADO

```
for i in range(500):
    tipo = random.choice(tipos)
    texto = random.choice(mensagens_exemplo[tipo])
    dados.append([texto, tipo])

df = pd.DataFrame(dados, columns=["mensagem", "categoria"])

print(df.head())

# =====
# 2. NLP - CLASSIFICAÇÃO AUTOMÁTICA
# =====

vectorizer = CountVectorizer()
X = vectorizer.fit_transform(df["mensagem"])
y = df["categoria"]

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)

modelo = MultinomialNB()
modelo.fit(X_train, y_train)

pred = modelo.predict(X_test)

print("\nRELATÓRIO DE CLASSIFICAÇÃO:")
print(classification_report(y_test, pred))

# =====
# 3. ESTATÍSTICA - ANÁLISE DOS ATENDIMENTOS
# =====

freq = df["categoria"].value_counts()
print("\nFrequência por categoria:")
print(freq)

freq.plot(kind="bar", title="Distribuição das categorias")
plt.show()

# =====
# 4. PESQUISA OPERACIONAL - SIMULAÇÃO DE FILA DE ATENDIMENTO
# =====
```

PROJETO INTEGRADO

```
# Simulação de tempos (em minutos)
tempo_chegada = np.random.exponential(scale=3, size=300)
tempo_atendimento = np.random.normal(loc=5, scale=1, size=300)

tempo_espera = np.maximum(tempo_atendimento - tempo_chegada, 0)

df_fila = pd.DataFrame({
    "chegada": tempo_chegada,
    "atendimento": tempo_atendimento,
    "espera": tempo_espera
})

plt.hist(df_fila["espera"], bins=20)
plt.title("Distribuição do Tempo de Espera")
plt.xlabel("Tempo (min)")
plt.show()

print("\nTempo médio de espera:", df_fila["espera"].mean())

# =====
# 5. SEGURANÇA — ANONIMIZAÇÃO DE TEXTO
# =====

def anonimizar(texto):
    return texto.replace("pedido", "[DADO]").replace("produto", "[DADO]")

df["mensagem_anonima"] = df["mensagem"].apply(anonimizar)
print("\nMensagens anonimizadas:")
print(df[["mensagem", "mensagem_anonima"]].head())
```

Perguntas sobre o código do COLAB:

A) NLP

1. Qual categoria teve melhor desempenho no classificador Naive Bayes?
2. Em suas palavras, como funciona a vetorização Bag of Words?

PROJETO INTEGRADO

B) Probabilidade e Estatística

3. Qual categoria teve maior frequência? O que isso indica?
4. Interprete o histograma de tempo de espera: é simétrico, assimétrico, disperso?

C) Pesquisa Operacional

5. O tempo médio de espera é adequado? Como poderia ser reduzido?
6. Cite duas ações que podem otimizar o fluxo de atendimento.

D) Segurança de Dados

7. O que é anonimização?
8. Quais dados sensíveis podem aparecer em mensagens de clientes?
9. Qual risco existe em armazenar textos sem anonimização?

Perguntas por disciplina

A) PROCESSAMENTO DE LINGUAGEM NATURAL (NLP)

1. Explique como o processo de pré-processamento textual (tokenização, normalização e remoção de ruído) pode influenciar o desempenho de um classificador de mensagens. Dê um exemplo prático.
2. Compare Bag of Words e TF-IDF. Em que situações TF-IDF produz resultados melhores e por quê?
3. O Naive Bayes assume independência entre as palavras. Em mensagens curtas de atendimento, essa suposição pode impactar negativamente os resultados? Justifique com base no contexto real do problema.

B) PROBABILIDADE E ESTATÍSTICA PARA ANÁLISE DE DADOS

4. Em um teste A/B aplicado ao tempo de resposta do atendimento, o grupo A teve média 6,2 min e o grupo B 5,1 min. Explique como você verificaria estatisticamente se essa diferença é significativa.

PROJETO INTEGRADO

5. Dado um histograma de tempo de espera assimétrico à direita, descreva qual medida de tendência central (média, mediana ou moda) é mais apropriada para representar esse conjunto. Justifique.

6. O tempo entre chegadas no call center segue uma distribuição exponencial. Explique o que isso significa em termos de probabilidade e por que esse modelo é amplamente usado em filas.

C) OTIMIZAÇÃO E PESQUISA OPERACIONAL

7. Explique como um modelo de simulação de eventos discretos poderia ajudar a decidir entre contratar mais atendentes ou investir em um chatbot inicial.

8. O tempo de chegada e atendimento no sistema foram modelados por distribuições estatísticas. Justifique por que essa modelagem é essencial para simulação e otimize uma sugestão de melhoria no processo.

D) SEGURANÇA DE DADOS

10. Mensagens de atendimento podem conter dados pessoais e até dados sensíveis. Explique como identificar esses elementos automaticamente e os riscos de mantê-los em texto puro.

11. Diferencie pseudonimização de anonimização e explique qual delas é mais adequada para sistemas de chatbot.

12. Em um pipeline de NLP, em que etapa a criptografia deve ser aplicada? Explique sua resposta com foco em segurança, desempenho e arquitetura.

NORMAS PARA ELABORAÇÃO E ENTREGA DA ATIVIDADE

O trabalho final deverá ser realizado no arquivo denominado Modelo, disponibilizado no AVA, obedecendo os critérios a seguir:

- a) Atenção aos prazos de postagens, acompanhe o cronograma das datas de postagem e correção das atividades em seu AVA.
- b) **A postagem** no AVA deve ser **em um único arquivo**, em formato (Word ou pdf), com tamanho máximo de 10MB, conforme Modelo. O sistema irá disponibilizar para correção apenas o último arquivo postado.
- c) **O trabalho deverá ser realizado individualmente.**

PROJETO INTEGRADO

- d) O trabalho final deve ser original e, portanto, não poderá haver trabalhos idênticos aos de outros grupos ou com reprodução de materiais extraídos da internet. Os trabalhos plagiados serão invalidados, sendo os alunos reprovados na atividade. Observe que a prática do plágio constitui crime, com pena prevista em lei (**Lei n.º 9.610**), e deve ser evitada no âmbito acadêmico.
- e) Desenvolver o projeto integrado a partir do template disponibilizado no AVA. Para nortear o desenvolvimento do projeto integrado que está sendo proposto, é necessário que sejam apresentados os resultados no template, seguindo as normas da ABNT.
- f) Em caso de dúvida para elaboração do trabalho, você deverá buscar orientações com seu tutor a distância.

O Projeto Integrado é sua chance de aplicar a teoria, simular desafios reais e desenvolver habilidades essenciais (solução de problemas, colaboração e pensamento crítico).

Dedique-se, sua construção de conhecimento hoje é a chave para o sucesso profissional amanhã.

CRITÉRIOS DE CORREÇÃO

CRITÉRIOS	PESO
Execução da Proposta <ul style="list-style-type: none">✓ Atendimento a todas as ações solicitadas na atividade.✓ Cumprimento do objetivo solicitado na proposta.✓ Presença de introdução, desenvolvimento e considerações finais, conforme indicado na proposta.✓ Qualidade técnica das imagens apresentadas, com atenção à resolução e nitidez, quando se aplicar.	60%

PROJETO INTEGRADO

Linguagem e Comunicação <ul style="list-style-type: none">✓ Clareza na comunicação das ideias.✓ Norma padrão, ortografia, aspectos redacionais, coesão e coerência.	30%
Normalização <ul style="list-style-type: none">✓ Presença de citações de autores e obras no decorrer do trabalho (quando se aplicar).✓ Inclusão adequada das referências bibliográficas utilizadas.	10%
TOTAL	100%

*Desejamos uma ótima atividade!
Equipe de professores*