

UNIVERSIDADE VIRTUAL DO ESTADO DE SÃO PAULO

Fábio Wakisaka
1800221

Modelo Preditivo de Turnover em Empresas

Vídeo de apresentação do Projeto Integrador

<<https://youtu.be/ETTbESwZ2V4>>

Americana - SP
2023

UNIVERSIDADE VIRTUAL DO ESTADO DE SÃO PAULO

Modelo Preditivo de Turnover em Empresas

Relatório Técnico-Científico apresentado na disciplina de Projeto Integrador para o curso de Engenharia de Computação da Universidade Virtual do Estado de São Paulo (UNIVESP).

Americana - SP
2023

WAKISAKA, Fábio. **Modelo Preditivo de Turnover em Empresas**. Relatório Técnico-Científico. Engenharia de Computação – **Universidade Virtual do Estado de São Paulo**. Polo Americana, 2023.

RESUMO

Organizações de portes diversos têm encontrado dificuldades no que se refere à retenção de talentos para o mercado, tendo em vista a elevada rotatividade. Alguns fatores mostram-se determinantes para tal rotatividade, e devido à sua complexidade, nem sempre pode ser prevista. A análise da evasão de funcionários é calculada através do *Turnover*. Neste projeto, será proposto um modelo preditivo baseado em aprendizado de máquina, que permita estimar o *Turnover* em pequenas organizações, auxiliando na tomada de decisões no gerenciamento de pessoas.

PALAVRAS-CHAVE: *Turnover*; Rotatividade; Aprendizado de máquina.

LISTA DE ILUSTRAÇÕES

Figura 1: Base de dados 'dataset.csv'	9
Figura 2: Distribuição de salários	10
Figura 3: Distribuição da Satisfação no Trabalho	10
Figura 4: Distribuição de Cargos	11
Figura 5: Boxplot Satisfação no Trabalho x Cargos.....	11
Figura 6: Admissões e demissões.....	12
Figura 7: Admissões e demissões por ano, Turnover	15
Figura 8: Pré-processamento da base de dados	15
Figura 9: Turnover	16
Figura 10: Modelo <i>Logist Regression</i>	16
Figura 11: Taxa de Turnover	17
Figura 12: Matriz de Confusão	18
Figura 13: Comparação <i>Turnover</i> real e previsto	19

SUMÁRIO

1 INTRODUÇÃO	1
2 DESENVOLVIMENTO	2
2.1 Objetivos	2
2.2 Justificativa e delimitação do problema	2
2.3 Fundamentação teórica	3
2.4 Aplicação das Disciplinas estudadas no Projeto Integrador	4
2.5 Metodologia	5
3 RESULTADOS	6
3.1 Solução Inicial	6
3.2 Solução Final	14
4. CONSIDERAÇÕES FINAIS	20
REFERÊNCIAS	21

1 Introdução

Turnover (ZENDESK, 2022), ou taxa de saída de funcionários, refere-se à dinâmica natural de rotatividade de funcionários em uma organização. Fatores diversos estão diretamente ligados a essa rotatividade. Em toda organização ou empresa percebe-se essa movimentação, tanto pelo ingresso de novos funcionários quanto pelos desligamentos através do processo demissional.

Mensurar o *Turnover* mostra-se importante devido aos custos inerentes do processo. Em relação à demissão, podemos citar os riscos e custos trabalhistas, pagamento de verbas rescisórias, custos já despendidos no treinamento do funcionário. Da mesma forma, a busca por profissionais no mercado demanda recursos consideráveis, como despesas no processo de seleção, tempo improdutivo com a vaga ociosa, dificuldade em se encontrar profissionais qualificados com salários condizentes à realidade da empresa.

Neste cenário, verifica-se que a retenção de talentos mostra-se mais vantajosa em relação ao processo de ingresso de novos funcionários e ao processo demissional.

De forma a auxiliar essa análise, deve-se dar maior atenção às ferramentas e análises de *Turnover*, de forma a se antever cenários desfavoráveis.

Neste Projeto Integrador, através de um modelo preditivo baseado em aprendizado de máquina, espera-se que seja possível estimar o *Turnover*. De forma simples, este projeto irá propor a análise de uma base de dados, e calcular a probabilidade de ingresso e saída de funcionários, computacionalmente.

2 Desenvolvimento

2.1 Objetivos

Este Projeto Integrador tem como objetivo desenvolver um modelo computacional, de aprendizado de máquina, que permita analisar uma base de dados de funcionários, e estimar probabilidades de *Turnover*.

Desenvolvido em *Python*, no ambiente de desenvolvimento *Jupyter*, o modelo irá abranger as seguintes etapas:

- Exploratória: modelar uma base de dados que permita extrair informações pertinentes para a aplicabilidade do modelo, bem como identificar técnicas de análise de dados a serem utilizadas;
- Descritiva: prototipar e adequar o modelo preditivo à base de dados, analisando parâmetros e refinando o projeto;
- Explicativa: consolidar os resultados obtidos, e verificar a real aplicabilidade do modelo.

2.2 Justificativa e delimitação do problema

Sustentabilidade, de forma ampla, abrange diversos setores da sociedade, com o objetivo final de se envidar esforços evitando-se desperdício, bem como a economia de recursos.

O despertar da consciência em relação à necessidade de se otimizar recursos e processos, ainda que amplamente divulgado, comumente priorizando a questão ambiental, pode e deve ser estendido para outros setores. Por sustentabilidade, entende-se não apenas a necessidade de conscientização global, como a preservação dos recursos naturais do planeta, mas também com iniciativas pontuais, que também se mostrem efetivas, sendo parte de um todo.

É fato que, com a recente emergência da pauta ESG (TOTVS, 2022), ou Governança Ambiental, Social e Corporativa, do inglês *Environmental, Social, and Corporate Governance*, se mostram imprescindíveis os aprimoramentos em processos, nos ambientes empresariais, que permitam de fato consolidar os objetivos propostos.

A sustentabilidade empresarial, seu sucesso, e efetiva consecução do pressuposto em sua política de governança ESG, notadamente requer o esforço no tratamento de seus recursos, em especial as pessoas.

Uma efetiva gestão de recursos humanos, também de forma analítica, com o auxílio de novas tecnologias, mostra-se dessa forma fundamental para este novo cenário de entregas sustentáveis.

2. 3 Fundamentação teórica

Um ambiente de trabalho adequado em uma empresa pode ser caracterizado pela baixa rotatividade de seus funcionários. Empresas com elevadas taxas de demissão, baixo tempo de permanência no emprego, processos seletivos dispersos ou incoerentes, acabam por refletir uma má gestão em recursos humanos. Ainda, a própria subjetividade (SOUZA) na relação entre o funcionário e sua empresa demanda profissionais devidamente capacitados para os processos admissionais, fato este nem sempre levado em consideração, notadamente em empresas de menor porte.

O objetivo fim de uma empresa só pode ser alcançado através de um corpo de funcionários alinhados à sua estratégia corporativa. Assim, a prospecção desses funcionários deve ser realizada de forma objetiva, envidando-se esforços para a correta identificação de talentos, sob risco de perdas financeiras, por vezes sucessivas, pela má gestão nos processos seletivos.

Tanto a identificação quanto a retenção desses talentos se mostra um trabalho complexo, devido à alta competitividade entre as empresas. Uma seleção imprecisa também impede a correta identificação de recursos que venham, de fato, agregar à empresa. Ainda, a condução de todo o ciclo de permanência do funcionário requer uma gestão eficiente, que atente-se à sua satisfação, ao seu retorno, proporcionando-lhe condições adequadas para sua retenção e minimização de eventuais descontentamentos.

Além disso, a retenção de talentos é importante para garantir que os custos com capital humano e a produtividade da empresa não sejam impactados pelo turnover. As estratégias de retenção garantem que o time de colaboradores se mantenha satisfeito e engajado, aumentando sua produtividade e ajudando a organização a crescer. (CRUZ, 2022)

Verifica-se, dessa forma, a importância de se analisar indicadores como o *Turnover*.

2.4 Aplicação das Disciplinas estudadas no Projeto Integrador

Considera-se para este projeto a utilização de forma prática de disciplinas como Programação de Computadores, a qual será utilizada para o desenvolvimento do protótipo do projeto.

Retomam-se ainda conhecimentos das disciplinas de Administração e Economia, fontes das quais serão utilizados para contextualização aderente ao tema norteador 'Sustentabilidade'.

2.5 Metodologia

A metodologia utilizada foi baseada em *Design Thinking* (FIA, 2018).

Inicialmente, contextualizou-se o problema identificado em sua fase de entendimento e compreensão. Trata-se da **Imersão** na questão de rotatividade de pessoal, o qual requer um tratamento analítico, considerando-se eventuais vieses nas análises subjetivas.

Foi então realizada a fase de **Análise e Síntese**, a fim de se gerar a base de dados, permitindo assim o esboço do projeto em *Python*, extraíndo-se *insights* e iniciando um modelo que se mostre, desde o início, funcional

O esboço funcional envolve as fases de **Ideação e Prototipação** em *Design Thinking*. Foram realizados testes contínuos, aplicados à base de dados, permitindo o refinamento até que o protótipo entregue um mínimo de informações condizentes ao problema proposto.

Finalmente, em sua fase de **Implementação Final**, o projeto permitiu, de fato, a identificação de dados pertinentes sobre o *Turnover* para a base de dados utilizada.

3 Resultados

3.1 Solução Inicial

Para a análise do *Turnover*, é necessária uma base de dados que possibilite, pelo menos, informações referentes ao desligamento de funcionários em uma empresa.

$$\text{Turnover} = \frac{(\text{Número de funcionários que deixaram a empresa durante o período})}{(\text{Número médio de funcionários durante o período})} * 100$$

Pressupondo que no início do ano a empresa possuía 100 funcionários, ao final do mesmo ano ela passou a contar com 110 colaboradores, tendo havido a saída de 40 profissionais durante esse período. Nesse caso, o cálculo do turnover seria:

$$\text{Turnover} = (40 / 100) * 100$$

$$\text{Turnover} = 40\%$$

Nesse caso, o *Turnover* da empresa foi de 40%. Isso significa que, em média, a cada 100 funcionários que a empresa tinha no início do ano, 40 deixaram a empresa ao longo do mesmo ano.

Percebe-se que o *Turnover* é elevado. Pode-se pressupor que há algo no ambiente corporativo, que esteja levando os colaboradores a sair da empresa.

São diversos os fatores que impactam diretamente no desligamento de funcionários, como por exemplo:

Fatores internos:

- Falta de perspectiva de crescimento profissional dentro da empresa;
- Falta de reconhecimento e valorização do trabalho desempenhado;
- Condições inadequadas de trabalho, como excesso de carga horária, falta de equipamentos adequados, clima organizacional desfavorável, entre outros;

- Falta de treinamento e capacitação adequados;
- Salários e benefícios insuficientes ou abaixo da média de mercado;
- Conflitos com colegas de trabalho ou gestores.

Fatores externos:

- Ofertas de trabalho mais atraentes em outras empresas;
- Mudança de cidade ou país, geralmente motivada por questões pessoais, como casamento, filhos, necessidade de cuidar de familiares, entre outras;
- Crises econômicas que podem levar as empresas a reduzirem seus quadros de funcionários;
- Problemas de saúde que impedem o funcionário de continuar trabalhando;
- Mudanças na legislação trabalhista que podem afetar os benefícios e direitos dos funcionários.

Pode-se tentar estimar a quantidade de funcionários que vão se desligar da empresa, analisando-se as diversas variáveis. Notadamente, essas variáveis devem estar sendo coletadas de forma consistente pela empresa, para que, em conjunto com ações de retenção interna, seja possível avaliar a efetividade da redução do *Turnover*.

Como exemplo, podemos citar:

- Oferecer programas de treinamento e desenvolvimento de carreira para os funcionários crescerem profissionalmente dentro da empresa;
- Valorizar o trabalho dos funcionários com programas de incentivo, remuneração justa e feedback constante;
- Melhorar as condições de trabalho, oferecendo equipamentos adequados e promovendo um clima organizacional positivo;
- Oferecer benefícios e salários compatíveis com o mercado e avaliá-los regularmente;

- Promover um ambiente saudável no trabalho com políticas claras de relacionamento e resolução de conflitos;
- Estabelecer um plano de carreira para acompanhar o desenvolvimento dos funcionários;
- Promover a qualidade de vida dos funcionários com flexibilidade no horário de trabalho e iniciativas de bem-estar.

Como forma de se tentar analisar o *Turnover*, decidiu-se pela geração de uma base de dados, fictícia, que permita a extração de informações relevantes e condizentes com um cenário de descontentamento em uma empresa. Essa base de dados foi gerada de forma computacional em *Python*, e apresenta os seguintes campos:

- nomeCompleto: nome completo do funcionário
- sexo: sexo do funcionário (masculino ou feminino)
- idade: idade do funcionário
- cargo: cargo do funcionário (Vendedor, Supervisor, Gerente ou Superintendente)
- salario: salário do funcionário
- satisfacaoTrabalho: nível de satisfação do funcionário no trabalho (0 a 10)
- dataAdmissao: data de admissão do funcionário (formato YYYY-MM-DD)
- dataDesligamento: data de desligamento do funcionário, se houver (formato YYYY-MM-DD).

```
dataset.csv X
C: > Users > vani_ > dataset > dataset.csv
1 nomeCompleto,sexo,idade,cargo,salario,satisfacaoTrabalho,dataAdmissao,dataDesligamento
2 Gabrielly Melo,masculino,21,Vendedor,3776.63,10,2019-04-28,2020-02-23
3 Srta. Fernanda Porto,masculino,41,Superintendente,9797.26,10,2017-07-27,2020-01-18
4 Dr. André da Mata,masculino,68,Superintendente,9232.15,10,2017-12-15,2020-02-04
5 Maitê Pires,feminino,57,Gerente,6058.31,8,2015-04-05,2015-07-31
6 Luana da Conceição,masculino,40,Superintendente,8529.22,10,2014-05-16,2016-04-05
7 Pedro Lucas Caldeira,feminino,64,Superintendente,11774.35,8,2010-06-27,2012-09-01
8 Dr. Lorenzo Almeida,feminino,36,Superintendente,8704.18,8,2011-06-23,2022-09-19
9 Marcos Vinicius Carvalho,feminino,28,Gerente,6586.83,10,2019-08-19,2022-04-23
10 Ana Livia da Cruz,feminino,36,Gerente,6882.99,8,2019-08-19,2020-05-13
11 Isabel Monteiro,masculino,65,Supervisor,3042.21,10,2012-09-03,2018-06-12
12 André Ferreira,feminino,48,Vendedor,3470.76,8,2018-06-10,2019-05-01
13 Maria Clara Jesus,masculino,25,Vendedor,2486.63,9,2021-08-21,2022-04-07
14 Pietra Costa,feminino,33,Vendedor,3071.05,8,2012-05-18,2018-04-18
15 Luiz Henrique Mendes,feminino,32,Vendedor,3121.01,9,2020-07-19,2022-07-31
16 Dra. Ana Clara Cardoso,masculino,45,Vendedor,2794.89,8,2019-06-30,2020-04-12
17 Dr. Thomas Porto,feminino,60,Vendedor,2888.33,10,2022-12-25,2023-01-26
18 Sofia Gonçalves,feminino,32,Supervisor,4213.97,8,2011-11-26,2012-09-26
19 Clara Ramos,feminino,28,Superintendente,9857.1,10,2021-09-05,2023-01-17
20 Dra. Ana Vitória Vieira,masculino,20,Vendedor,3224.22,10,2010-06-23,2020-05-15
21 Agatha Novaes,feminino,19,Gerente,6343.72,8,2022-07-23,2022-12-14
22 Pietra Ribeiro,masculino,20,Vendedor,2597.8,9,2020-09-23,2022-05-27
23 Miguel Caldeira,feminino,31,Gerente,5994.98,10,2016-11-14,2021-12-21
24 Stella Azevedo,masculino,28,Superintendente,8144.86,9,2022-11-05,2022-12-04
25 Cecília Fogaca,masculino,54,Gerente,5747.74,9,2011-02-28,2012-09-17
```

Figura 1: base de dados 'dataset.csv'

Nesta base de dados, foram definidas condições que a tornem altamente enviesada. Os registros gerados, dessa forma, tendem a refletir um ambiente de trabalho com indícios de insatisfação, a saber:

- os salários variam de acordo com o cargo:
 - Vendedores: entre R\$ 2.000,00 e R\$ 4.000,00
 - Supervisor: entre R\$ 3.000,00 e R\$ 5.000,00
 - Gerentes: entre R\$ 5.000,00 e R\$ 7.000,00
 - Superintendente: entre R\$ 8.000,00 e R\$ 12.000,00
- para Gerentes e Superintendentes, as notas de satisfação no trabalho são mais altas, entre 5 e 8;
- quando a data de admissão e de desligamento forem muito próximas, atribuiu-se como nota de satisfação no trabalho valores entre 0 e 4;
- 94% da base de dados está empregada (de 10.000 funcionários, 600 saíram);
- Vendedores são em maior número, com Gerentes, Supervisores e Superintendentes em números relativamente elevados para o mercado (distribuição de 7:2:1:1).

Segue abaixo a distribuição gerada pelo código:

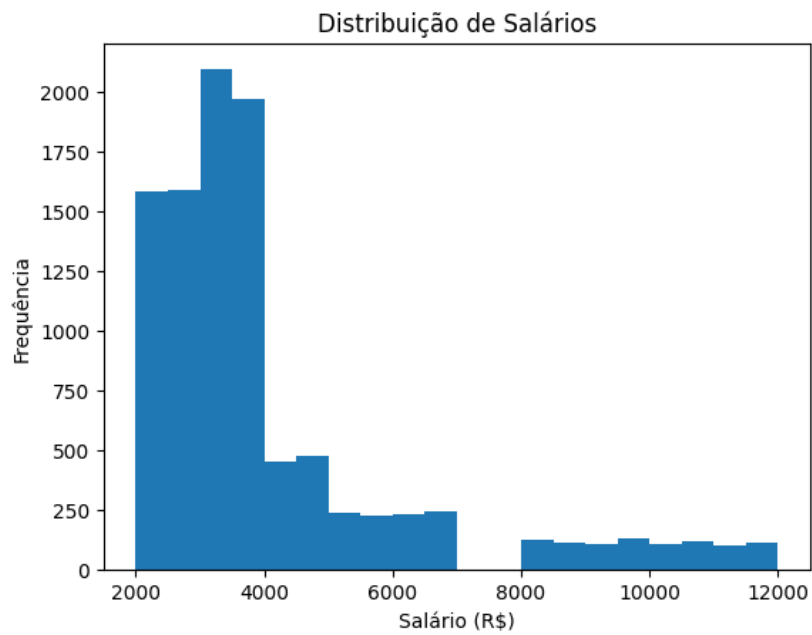


Figura 2: Distribuição de salários

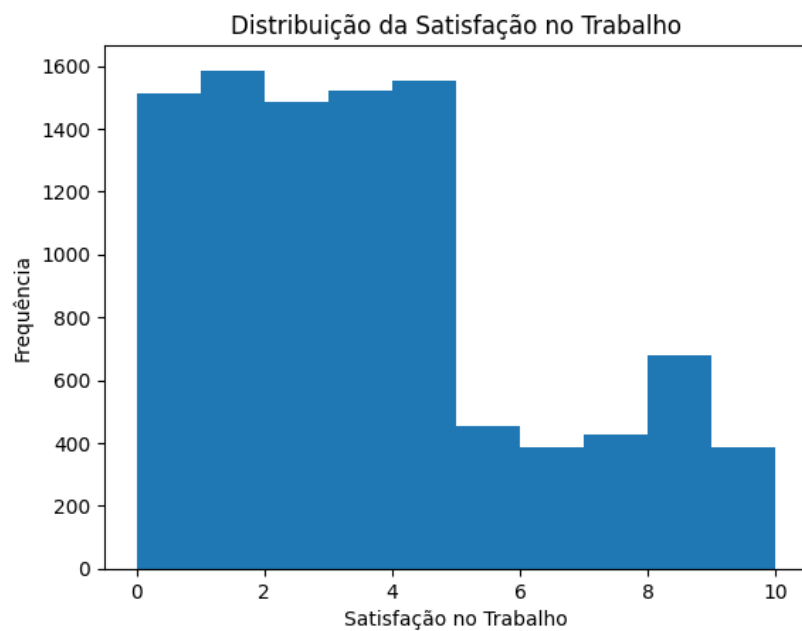


Figura 3: Distribuição da Satisfação no Trabalho

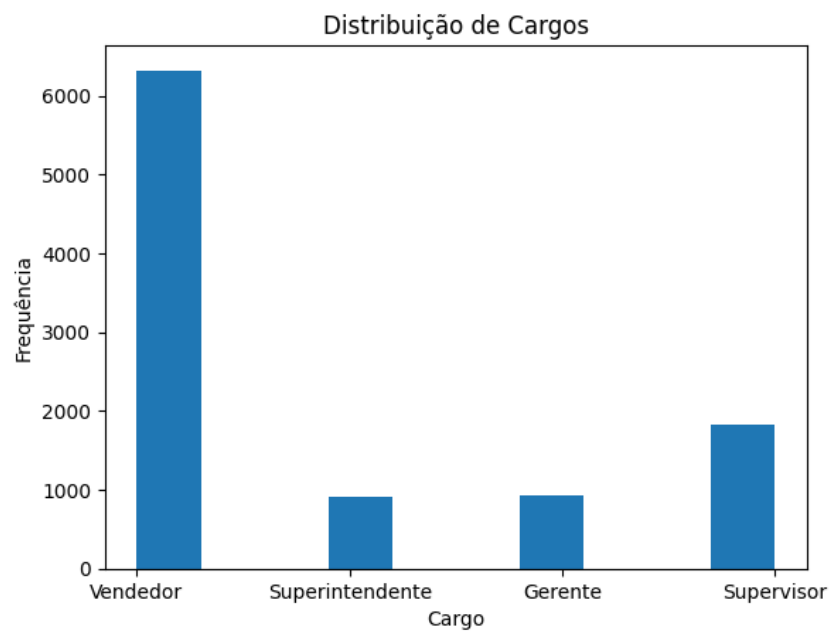


Figura 4: Distribuição de Cargos

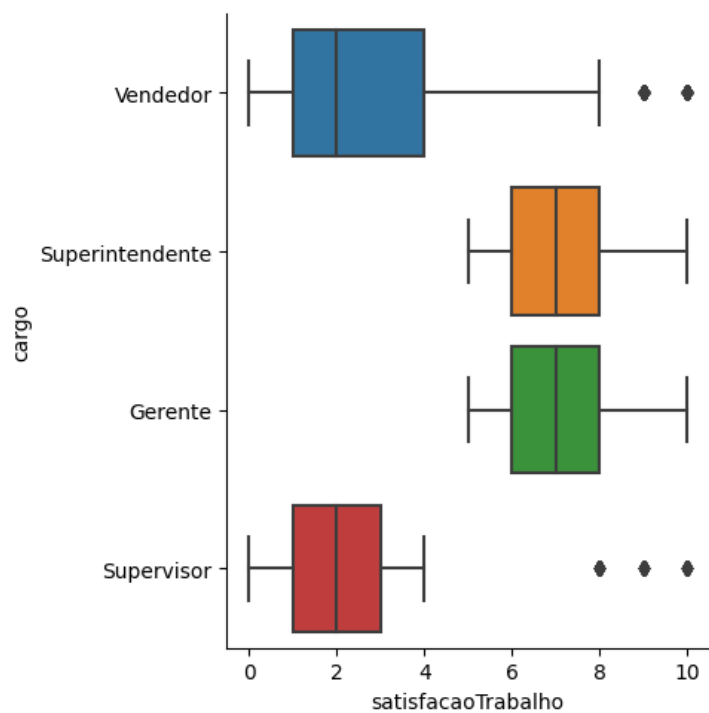


Figura 5: Boxplot Satisfação no Trabalho x Cargos

Basicamente, percebe-se que neste ambiente de trabalho há relativa insatisfação. Notadamente, quem ganha menos, ou tem cargo de vendedor, está menos satisfeito. Há também elevada pressão, pois há muitos superiores, em todos os cargos.

No gráfico abaixo, é possível verificar ao longo dos anos a quantidade de admissões e de desligamentos:

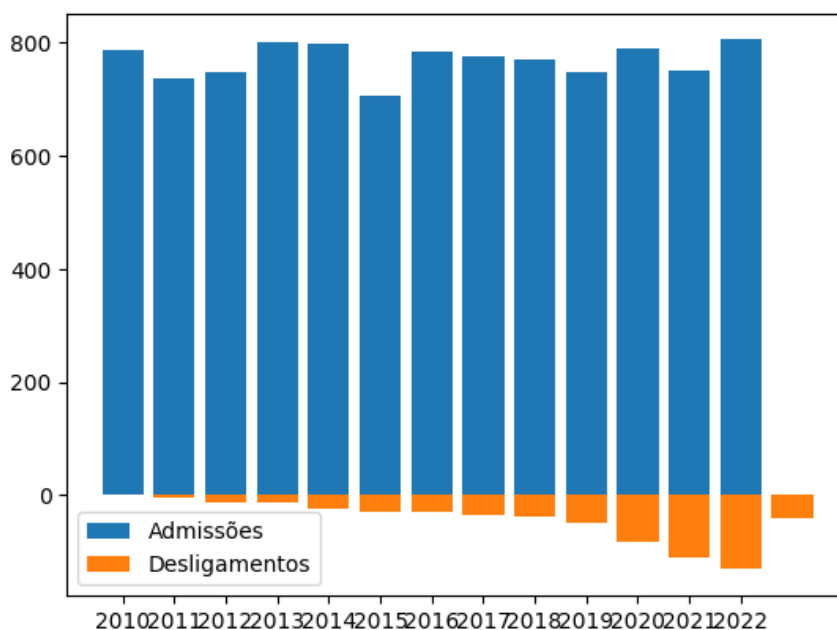


Figura 6: Admissões e Desligamentos

Em um mesmo gráfico, verificamos que a taxa de *Turnover* tem aumentado:

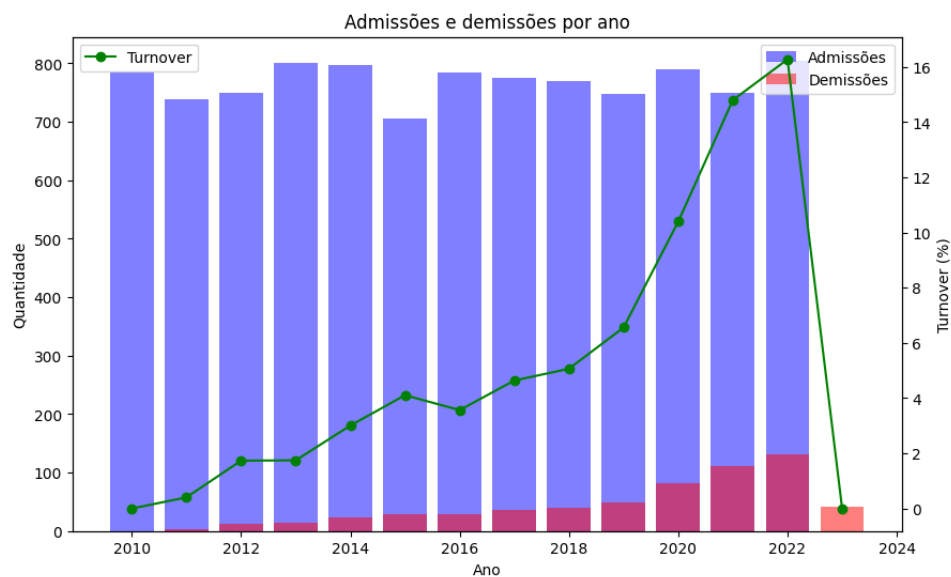


Figura 6: Admissões e demissões por ano, Turnover

É possível verificar, dessa forma, que são necessárias medidas sustentáveis, que revertam essa tendência de aumento nos desligamentos.

3.1 Solução Final

Considerando o histórico de admissões e desligamentos de uma base de dados (dataset), pode-se pressupor que alguns fatores estejam diretamente ligados à taxa de satisfação, a qual pode ser determinante para um pedido de permanência ou de demissão, por exemplo.

Para a base fictícia modelada, foi acrescentado o viés negativo, onde elevadas taxas de insatisfação condizem com o cenário de data de admissão e de desligamento muito próximas. Ou seja, se o funcionário permanece pouco tempo na empresa, ele possivelmente está insatisfeito.

Gênero e idade também podem ser fatores intrínsecos à insatisfação. Notadamente, considerando a diversidade de ambientes corporativos, bem como suas culturas corporativas, é possível que ainda existam problemas internos, talvez ainda não identificados, ou mesmo ainda não solucionados, que perpetuam esse problema. Entretanto, em nossa base de dados, tal viés não foi implementado de forma proposital, cabendo ao modelo computacional fazer sua avaliação, a seu critério.

O código fonte está disponível no repositório abaixo:

<https://github.com/fabiowakisaka/projetointegradorviii>

Trata-se de um modelo computacional em *Python*, que irá utilizar a técnica estatística de Regressão Logística (SANTOS, 2019). Tal modelo, para a base de dados criada, terá como objetivo determinar se um funcionário irá sair ou não da empresa (valor binário), utilizando-se os valores disponíveis para análise, e posteriormente calcular o *Turnover*.

Inicialmente, foi necessário tratar a base de dados, de forma que o tempo de trabalho dos funcionários pudesse ser calculado em dias. O gênero dos funcionários também foi atualizado, garantindo a possibilidade de análise em formatos numéricos, em substituição aos formatos categóricos.

```

In [254]: from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()
df['sexo'] = le.fit_transform(df['sexo'])

In [255]: cargo_dummies = pd.get_dummies(df['cargo'], prefix='cargo')
df = pd.concat([df, cargo_dummies], axis=1)
df.drop('cargo', axis=1, inplace=True)

In [256]: df['dataAdmissao'] = pd.to_datetime(df['dataAdmissao'], format='%Y-%m-%d')
df['dataDesligamento'] = pd.to_datetime(df['dataDesligamento'], format='%Y-%m-%d', errors='coerce')
df['tempo_de_trabalho'] = (df['dataDesligamento'] - df['dataAdmissao']).dt.days

```

Figura 7: Pré-processamento da base de dados

Criou-se a variável alvo 'saiu', que será utilizada para o treinamento do modelo. Outra atualização relevante foi a substituição de registros nulos para a quantidade de dias trabalhados, tendo em vista a base de dados conter funcionários ainda em atividade. Para tanto, foi atribuído um valor máximo, baseado na data atual, em substituição aos valores NaN.

```

In [271]: data_ficticia = pd.to_datetime('2023-03-25')
df['dataDesligamento'].fillna(data_ficticia, inplace=True)

# cria coluna alvo
data_atual = pd.to_datetime('today').date()
df['saiu'] = df['dataDesligamento'].apply(lambda x: 0 if pd.isnull(x) or x.date() == data_atual else 1)
# remove coluna dataDesligamento
df.drop('dataDesligamento', axis=1, inplace=True)

tempo_trabalho_max = max(df['tempo_de_trabalho'])
df['tempo_de_trabalho'].fillna(tempo_trabalho_max, inplace=True)

df

```

Out[271]:

id	sexo	idade	salario	satisfacaoTrabalho	dataAdmissao	cargo_Gerente	cargo_Superintendente	cargo_Supervisor	cargo_Vendedor	tempo_de_trabalho	saiu
elo	1	21	3776.63	10	2019-04-28	0	0	0	1	301.0	1
ida	1	41	9797.26	10	2017-07-27	0	1	0	0	905.0	1
da	1	68	9232.15	10	2017-12-15	0	1	0	0	781.0	1

Figura 8: Pré-processamento da base de dados

Realizado o pré-processamento de dados (*data preprocessing*), pode-se calcular o Turnover da base de dados como um todo, cujo resultado é condizente com o definido para a base criada:

```
In [259]: turnover_train = df[df.index.isin(X_train.index)]['saiu'].sum() / len(X_train)
print(f'Turnover da base de treinamento: {turnover_train:.2%}')
|
```

Turnover da base de treinamento: 5.90%

Figura 9: *Turnover*

Em Ciência de Dados, utiliza-se uma técnica para que seja possível treinar um modelo preditivo. Esse treinamento requer que a base de dados seja dividida em duas partes, de maneira que na parte chamada de treinamento, o modelo consiga extrair informações relevantes, que possam ser testadas na outra parte, a parte de testes. Comumente a proporção de divisão entre as partes fica entre 70/30 ou 80/20, percentualmente.

Realizada a divisão da base de dados em X_{train} e X_{test} , descartando-se os campos que não se mostrem relevantes para o modelo (nomeCompleto, e campos de datas já utilizados para o cálculo de dias), inicia-se o treinamento do modelo *Logistic Regression*, pertencente à biblioteca scikit-learn (SCIKIT-LEARN):

```
In [260]: from sklearn.model_selection import train_test_split

X = df.drop(['nomeCompleto', 'dataAdmissao'], axis=1)
y = df['saiu']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

In [261]: from sklearn.linear_model import LogisticRegression

logreg = LogisticRegression(max_iter=1000)
logreg.fit(X_train, y_train)

Out[261]: LogisticRegression
LogisticRegression(max_iter=1000)
```

Figura 10: Modelo *Logist Regression*

Como resultado do treinamento utilizando-se o modelo preditivo, é possível calcular o Turnover em ambas as bases, de treinamento e de testes, tendo em vista

já terem sido estimadas a probabilidade de desligamento de cada funcionário da base de testes:

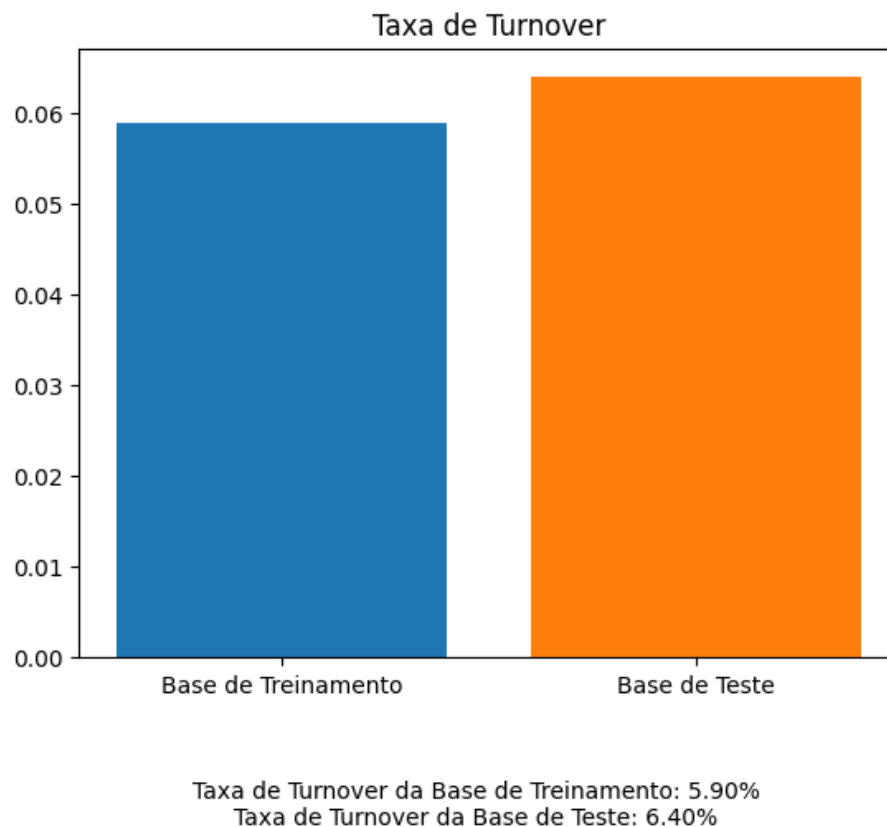


Figura 11: Taxa de Turnover

Pelo gráfico, é possível verificar que o modelo estimou, de forma aparentemente assertiva, a estimativa de Turnover para a base de testes. Para confirmar a eficácia do modelo, podemos utilizar a Matriz de Confusão:

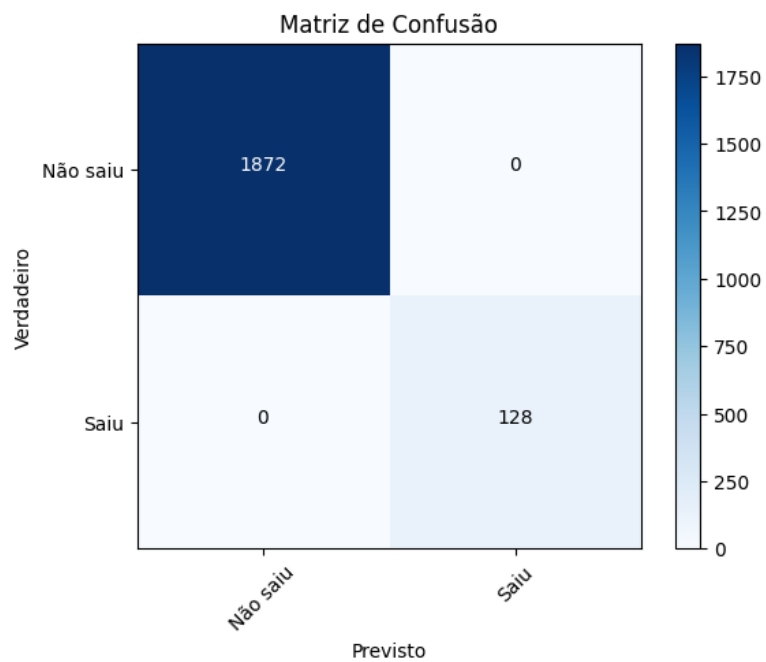


Figura 12: Matriz de Confusão

Pela análise da Matriz de Confusão, verifica-se que para os 2.000 (20%) registros da base de testes, foram previstos os 1872 Falsos-Verdadeiros (funcionários que não saíram) e os 128 Verdadeiros-Positivos (funcionários que saíram).

Finalmente, comparando o *Turnover* da base real de dados total, e o *Turnover* previsto para a base de testes, temos resultados próximos, validando o modelo treinado.

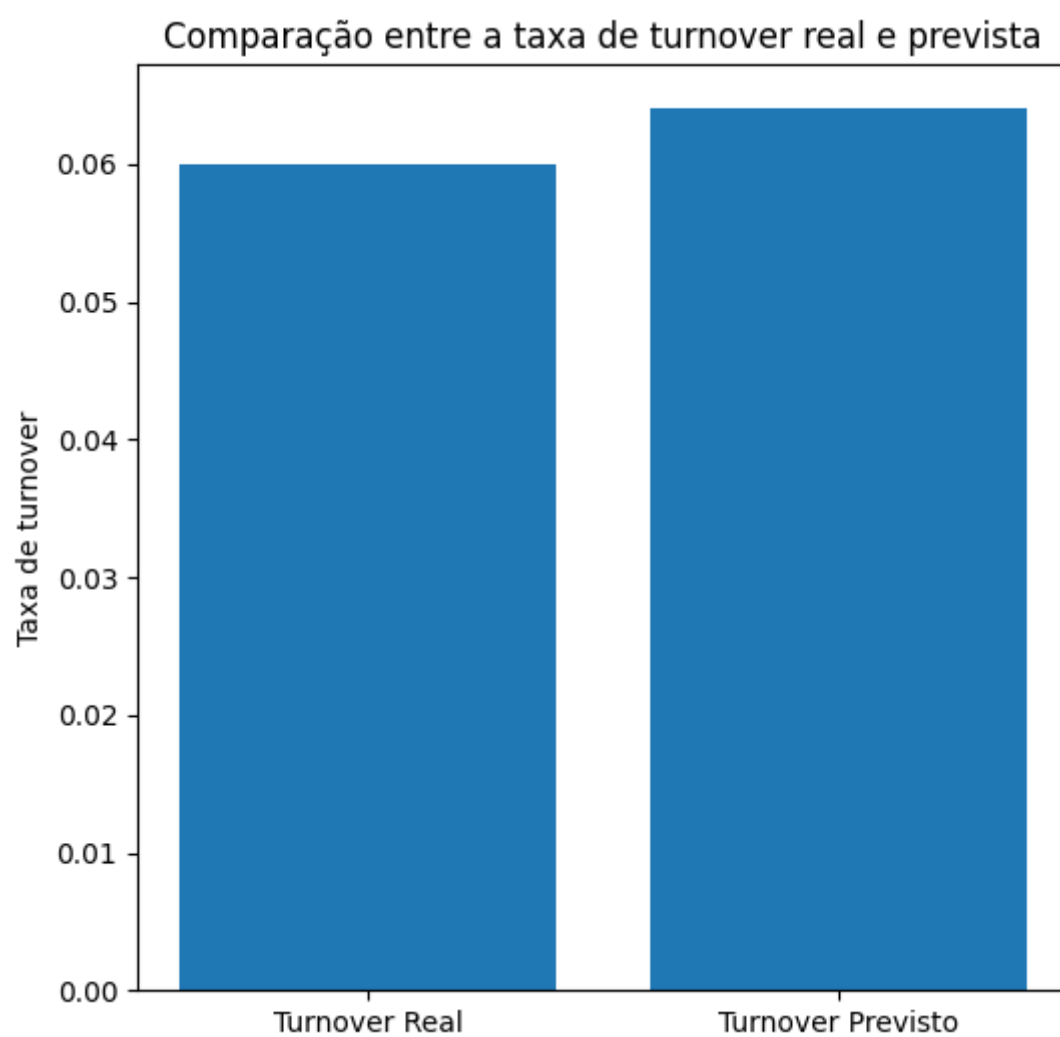


Figura 13: Comparação *Turnover* real e previsto

4 Considerações Finais

Ainda que focado na estimativa de *Turnover* para uma determinada base de dados, é importante compreender a análise computacional proposta neste projeto, a qual se mostra aderente à questão da sustentabilidade, em especial a sustentabilidade empresarial.

Em que pese a prevalência de direcionamentos específicos para a sustentabilidade, como exemplos, investimentos em pesquisas sobre sequestro de carbono, participação em fundos de pesquisa voltados para demandas ecológica e ambientais, substituição de copos descartáveis por canecas, é importante sempre estar atento ao recurso primordial no ambiente corporativo: cuidar da pessoas.

Práticas ESG têm como foco a geração de valor, tanto para clientes internos quanto externos. De forma que, não se utilizar de recursos tecnológicos internamente, no lidar com as pessoas, se mostra um relativo contrassenso.

A evasão de talentos, a dificuldade de retenção, o não aprimoramento de práticas corporativas que estejam voltadas para o bem estar e engajamento de seus funcionários, compromete a própria capacidade de sobrevivência de uma empresa.

Neste projeto, esperou-se ser possível criar um modelo preditivo simples, que pudesse estimar a taxa de evasão e rotatividade em empresas. Como resultado, foi possível verificar que, para variáveis simples, que possivelmente sejam relevantes para o desligamento de um funcionário, a previsibilidade pode ser modelada por análise de dados históricos.

De forma a se aprimorar o modelo, considera-se a possibilidade de utilização de outras técnicas de aprendizado e segmentação por agrupamentos, a serem implementados.

Referências

CRUZ, Leonardo. **Retenção de talentos: o que é, importância e como implementar.** **Creditas**, 9 de dez. de 2022. Disponível em <<https://www.creditas.com/rh-estrategico/retencao-de-talentos/>>. Acesso em 18 de mar. de 2023.

FIA. **Design Thinking: o que é, como aplicar e passo a passo.** **FIA Business School**, 18 de jul. de 2018. Disponível em <<https://fia.com.br/blog/design-thinking/>>. Acesso em 18 de mar. de 2023.

SANTOS, V. COMPUTER SCIENCE MASTER. **O que é regressão logística e como aplicá-la usando Python.** 16 de jan. de 2019. <<https://www.computersciencemaster.com.br/o-que-e-regressao-logistica-e-como-aplica-la-usando-python/>>. Acesso em 25 de mar. de 2023.

SCIKIT-LEARN. **sklearn.linear_model.LogisticRegression.** <https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html/>. Acesso em 25 de mar. de 2023.

SOUZA, A. R.. **CULTURA & GESTÃO. O indivíduo e sua subjetividade nas organizações.** Disponível em <<https://culturaegestao.com.br/blog/o-individuo-e-sua-subjetividade-nas-organizacoes>>. Acesso em 18 de mar. de 2023.

SOUZA, E. G.. **Entendendo o que é Matriz de Confusão com Python.** 27 de mar. de 2019. <<https://medium.com/data-hackers/entendendo-o-que-%C3%A9-matriz-de-confus%C3%A3o-com-python-114e683ec509/>>. Acesso em 18 de mar. de 2023.

TOTVS. **ESG: o que é, como funciona, vantagens e características.** **TOTVS**, 28 de out. de 2022. Disponível em <<https://www.totvs.com/blog/business-perfomance/esg/>>. Acesso em 18 de mar. de 2023.

ZENDESK. **O que é turnover? Tipos, causas e como calcular esse indicador.** **ZENDESK**, 02 de dez. de 2022. <<https://www.zendesk.com.br/blog/o-que-e-turnover/>>. Acesso em 18 de mar. de 2023.