

Inteligência Artificial no Processo de Seleção de Candidatos a Vagas de Emprego

1st Francisca Amanda Miranda de Paula

Centro Universitário IESB

Centro Universitário Instituto de Educação Superior de Brasília

Brasília, Brasil

Matrícula - 1812082052

2nd Matheus Henrique Santos Lucas

Centro Universitário IESB

Centro Universitário Instituto de Educação Superior de Brasília

Brasília, Brasil

Matrícula - 1822082032

Resumo O presente projeto consiste em desenvolver um aplicativo na linguagem flutter, que inclui técnicas de inteligência artificial para selecionar automaticamente o candidato que se encaixa melhor com a oportunidade de emprego em questão. Para o desenvolvimento da aplicação, com o objetivo de ter um projeto mais limpo, organizado e fácil de receber manutenção, foram implementados alguns recursos, sendo eles: arquitetura limpa (Clean Architecture), padrão de projeto MVVM, spring boot e lombok na API em Java etc.

Palavras-chave inteligência artificial, regressão linear, correlação linear.

Abstract—The present project consists in developing an application in flutter language, that includes techniques of artificial intelligence to automatically select the candidate who best fits the job opportunity. For application development, with the goal to have a cleaner project, more organized and easier to receive maintenance, was implemented some resources, like: clean architecture (Clean Architecture), MVVM design pattern, spring boot and lombok frameworks in the API in the Java etc.

Keywords—artificial intelligence, linear regression, linear correlation.

I. INTRODUÇÃO

O uso da inteligência artificial dentro da área de recursos humanos traz muitos benefícios ao agilizar o processo de contratação, além de torná-los mais eficazes. Entre as diversas técnicas que podem ser aplicadas para implementar esta funcionalidade, existe a regressão linear que aprende a prever o melhor perfil treinando por meio de dados históricos que já vem com os resultados pré-definidos.

II. CONTEXTUALIZAÇÃO

A candidatura às vagas de emprego geralmente é bastante concorrida pelo fato de ter muitos participantes para fazer a avaliação do perfil, tornando o processo de contratação lento na maioria das vezes. Sendo assim, é interessante agilizar o processo de seleção, visto que a demora pode prejudicar tanto a empresa quanto os interessados pela vaga.

Portanto, muitas empresas quando estão contratando, ficam com certas demandas paradas somente esperando o novo funcionário assumir o cargo para então dar andamento e prosseguimento às tarefas, sendo que até o novo funcionário se adaptar às atividades leva algum tempo, o que pode levar ao atraso do cumprimento das demandas e regressão na produtividade da equipe como um todo. Quanto ao candidato, este permanece disponível para o mercado de trabalho por mais tempo do que deveria, perdendo até outras oportunidades de emprego.

III. PROBLEMA

Definir a melhor estratégia e técnica da inteligência artificial para usá-la na pré-seleção de candidatos em uma vaga específica, além de decidir quais variáveis serão levadas em conta para análise e tomada de decisão, obtendo uma alta taxa de precisão. Neste contexto, como selecionar o melhor candidato com a menor taxa de erros possíveis, tendo em vista que a quantidade de parâmetros para medir as habilidades de um profissional, considerando também os requisitos, são praticamente infinitos?

IV. OBJETIVO GERAL

Desenvolver um aplicativo para dispositivos móveis com sistema operacional Android, que faz uso da inteligência artificial para realizar uma contratação inicial de candidatos em vagas de emprego. Assim, para resolver o problema exposto, será necessário utilizar um algoritmo que faz ligação de variáveis, para comparar os atributos do perfil candidato com os atributos de requisito das vagas. Portanto uma das técnicas estudada que atende melhor o objetivo final do projeto para se obter os resultados esperados, é o algoritmo de regressão linear da área de inteligência artificial.

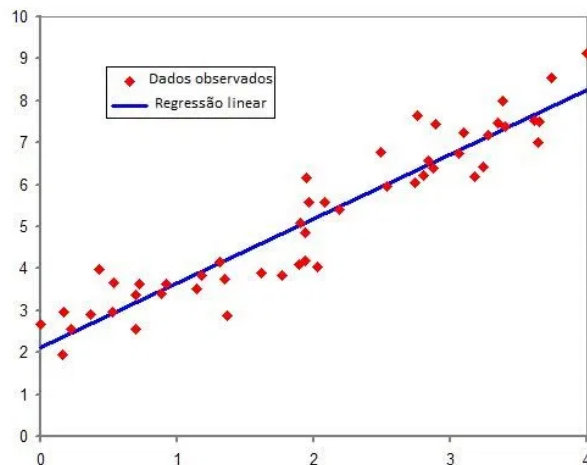
V. OBJETIVOS ESPECÍFICOS

- Desenvolver e consumir APIs;

- Usar padrões de projeto e boas práticas de programação, como os princípios S.O.L.I.D e Clean Architecture;
- Planejar e criar a interface do usuário;
- Integrar o backend com a Interface do usuário (frontend);
- Desenvolver funcionalidades para cadastro e autenticação de usuários;
- Conectar no banco de dados PostgreSQL;
- Subir aplicação para o Heroku;
- Aplicar a regressão linear para selecionar o candidato que é o mais apto para vaga de emprego;
- Incluir chat bot e mapa no projeto.

VI. REFERENCIAL TEÓRICO

No trabalho proposto, será estudado o funcionamento do Algoritmo de Regressão Linear dentro do contexto de recrutamento para vagas de emprego. Assim, como a técnica de análise de perfil para seleção envolve a ligação entre variáveis que existem uma boa correlação entre si, a regressão linear consegue verificar o quanto uma variável está relacionada à outra, por exemplo, comparando as habilidades do candidato com os requisitos das vagas de emprego, fazendo um ajuste linear nos dados obtidos em um diagrama de dispersão, conforme mostra o gráfico abaixo:



A. Regressão Linear

A regressão linear é um algoritmo supervisionado que aprende a fazer previsões a partir de análises de dados históricos que já tenham resultados pré-definidos, passando a obter um modelo de referência do que pode estar certo ou errado.

Para o procedimento da regressão linear aplicado na seleção de candidatos, é necessário utilizar um dataset que contenha dados históricos de um processo seletivo com a finalidade de treinar o algoritmo para que ele tenha a capacidade de prever os candidatos mais aptos para uma área específica. Para isso, será necessário definir quais serão as variáveis independentes, os atributos, que representam os requisitos e habilidades exigidas pela vaga em si.

Além disso, é necessário obter os valores da variável dependente, ou seja, o resultado da seleção individual de cada candidato. Por se tratar de resultados que possuem uma variável numérica, o método utilizado deverá ser o da regressão linear, pois para os casos em que o dado é categórico, utiliza-se a técnica de classificação linear.

Logo abaixo, tem-se um exemplo de um dataset voltado para vaga de Administrador de Banco de Dados (DBA):

	A	B	C	D	E	F	G	H	I	J	K	L	M	
ID	Inglês		Grau de Instrução	Gênero	Java	MySQL	PostgreSQL	Oracle	MongoDB	PHP	Git	Excel	Resultado	
1	Avançado		Pós-Graduação	F		8	7	8	10	3	10	7	10	242
2	Básico		Graduação	M		10	8	9	9	5	2	4	5	223
3	Intermediário		Graduação	F		5	5	3	7	7	8	5	6	145
4	Avançado		Graduação	F		2	4	10	9	8	9	7	9	239
5	Avançado		Mestrado	M		10	9	8	8	6	6	5	8	240
6	Básico		Graduação	M		6	4	7	6	7	5	8	10	180
7	Avançado		Pós-Graduação	F		7	6	5	5	8	10	9	7	188
8	Intermediário		Graduação	M		8	5	6	4	5	6	8	7	163

Para a regressão linear, todos os valores dos atributos deverão ser numéricos, pois é necessário realizar o cálculo da correlação linear de modo a medir o quanto aquela habilidade influencia nos resultados. Este cálculo da correlação, serve para selecionar quais variáveis independentes devem ser levadas em consideração no teste e treinamento do algoritmo, de modo a minimizar as diferenças entre o resultado esperado e o obtido.

Após converter os dados categóricos em numéricos, tem-se a seguinte tabela:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	ID	Inglês	Grau de Instrução	Gênero	Java	MySQL	PostgreSQL	Oracle	MongoDB	PHP	Git	Excel	Resultado
2	1	10	2	1	8	7	8	10	3	10	7	10	242
3	2	3	1	2	10	8	9	9	5	2	4	5	223
4	3	6	1	1	5	5	3	7	7	8	5	6	145
5	4	10	3	1	2	4	10	9	8	9	7	9	239
6	5	10	1	2	10	9	8	8	6	6	5	8	240
7	6	3	1	2	6	4	7	6	7	5	8	10	180
8	7	10	2	1	7	6	5	5	8	10	9	7	188
9	8	3	1	2	8	5	6	4	5	6	8	7	163

O próximo passo é calcular o valor da correlação cujo resultado varia entre -1 e 1. Assim, quanto mais próximo de 1, maior é a ligação da variável independente com a dependente (resultado), sendo congruente selecionar os requisitos que possuem estas características. Neste contexto, a correlação é classificada da seguinte forma:

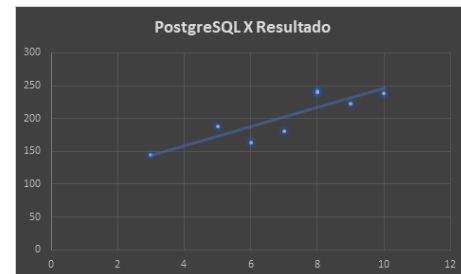
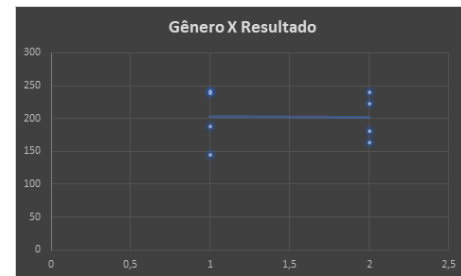
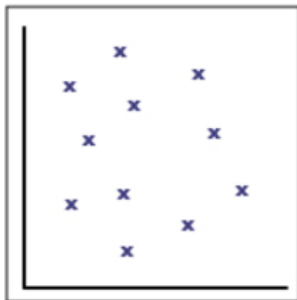
- Correlação Positiva: valor igual ou próximo de 1.



- Correlação Negativa: valor igual ou próximo de -1.



- Sem correlação: o resultado é igual a 0.



Após a obtenção dos resultados da correlação linear, foram excluídos os atributos que tinha valores próximos ou abaixo de 0, restando para análise as seguintes habilidades:

Para os testes e treinamentos do algoritmo, serão levados em conta os atributos com correlação linear que são mais próximos de 1.

ID	Inglês	Grau de Instrução	MySQL	PostgreSQL	Oracle	Excel	Resultado
1	10	2	7	8	10	10	242
2	3	1	8	9	9	5	223
3	6	1	5	3	7	6	145
4	10	3	4	10	9	9	239
5	10	1	9	8	8	8	240
6	3	1	4	7	6	10	180
7	10	2	6	5	5	7	188
8	3	1	5	6	4	7	163

Utilizando a função “correl” fornecida pelo Excel para calcular a correlação linear entre cada variável independente com a dependente, foram encontrados os seguintes valores:

Variável Independente	Correlação
Inglês	0,556634432
Grau de Instrução	0,482764478
Gênero	-0,027866615
Java	0,179739668
MySQL	0,526908128
PostgreSQL	0,863721481
Oracle	0,772389972
MongoDB	-0,292189954
PHP	0,04692566
Git	-0,236983287
Excel	0,36176296

Para obter o valor de cada resultado, foram utilizados diferentes pesos em cada habilidade, sendo estes escolhidos pelos recrutadores, considerando o nível de relevância de cada requisito existente para a vaga. Para o caso de uma das habilidades do candidato não ser um dos requisitos da vaga, neste será atribuído peso igual a zero. Sendo assim, o procedimento para chegar até o resultado, consiste em multiplicar cada pontuação que está na tabela pelo peso da respectiva habilidade, somando tudo no final. A seguir, tem-se um exemplo para calcular o resultado para o candidato de ID igual a 1 utilizando como exemplo, o dataset abaixo:

ID	Inglês	MySQL	PostgreSQL	Oracle	MongoDB	Git	Excel	Resultado
1	10	7	8	10	3	7	10	230
2	3	8	9	9	5	4	5	196
3	5	5	3	7	7	5	6	153
4	9	4	10	9	8	7	9	225
5	8	9	8	8	6	5	8	216
6	4	4	7	6	7	8	10	168
7	10	6	5	5	8	9	7	185
8	6	5	6	4	5	8	7	154

Abaixo, tem dois exemplos do diagrama de dispersão, sendo o primeiro uma correlação negativa e o segundo uma correlação positiva:

Habilidade	Peso
Oracle	7
PostgreSQL	6
MySQL	5
Inglês	4
Git	3
MongoDB	2
Excel	1

$$(4 \times 10) + (5 \times 7) + (6 \times 8) + (7 \times 10) + (2 \times 3) + (3 \times 7) + (1 \times 10) = 230$$

O tipo de regressão linear abordada neste trabalho é a múltipla, pelo fato de possuir mais de uma variável independente.

TRABALHOS CORRELATOS

Entre as inúmeras tecnologias existentes que automatizam processos, existem as que auxiliam no processo de contratação de candidatos às vagas de emprego. Como por exemplo, tem-se a ferramenta HireVue que é utilizada nas entrevistas online, sendo capaz de interpretar a análise da linguagem corporal, expressões faciais e tom de voz. Como estratégia de avaliação para decidir se o candidato é bom ou ruim, o entrevistado é comparado com os melhores funcionários da empresa, e a partir disso, selecionam os melhores para os recrutadores.

Entre as inúmeras tecnologias existentes que automatizam processos, existem as que auxiliam no processo de contratação de candidatos às vagas de emprego. Como por exemplo, tem-se a ferramenta HireVue que é utilizada nas entrevistas online, sendo capaz de interpretar a análise da linguagem corporal, expressões faciais e tom de voz. Como estratégia de avaliação para decidir se o candidato é bom ou ruim, o entrevistado é comparado com os melhores funcionários da empresa, e a partir disso, selecionam os melhores para os recrutadores.

A técnica conhecida como Lógica Fuzzy representa muito bem as incertezas e questões nebulosas que envolvem o contexto real no processo de seleção de um candidato para uma determinada vaga. Uma das formas de representar essa incerteza em conjunto com a Lógica Fuzzy, é utilizando o grau de pertinência para definir o resultado, ou seja, ao invés de designar que um perfil é bom ou ruim, a inteligência artificial quantifica esses atributos, transformando-os em um intervalo de valores, que varia entre 0 e 1, tornando possível inferir que um perfil pode ser 0.8 bom e 0.5 ruim, por exemplo.

O uso de Redes Neurais Artificiais (RNA) auxilia o RH pelo motivo de ter a capacidade de prever os candidatos que seriam os mais adequados às vagas, além de ter uma enorme habilidade em aprender no ambiente em que está inserido. Assim, o RNA, com o aumento das informações que adquire ao longo das avaliações dos currículos dos candidatos, cada vez mais a sua escolha se tornará mais assertiva, pois ele aprende com base na experiência.

Já o método denominado como Sistema Especialista, consiste em um sistema que se comporta como um ser humano

experiente em uma determinada área de atuação, podendo ser utilizado para fazer o papel de um especialista em RH. No Sistema Especialista, inicialmente é feita a coleta de dados a partir de um profissional experiente na sua área de atuação para que ele possa transmitir os seus conhecimentos para o sistema especialista. Logo em seguida, é criada uma base de conhecimento ou regras e, a partir disso, ela consegue fazer a inferência e dedução de novos conhecimentos sendo utilizada para a tomada de decisão, como decidir se vai selecionar ou não um candidato à vaga.

METODOLOGIA

A implementação do projeto é dividida em duas partes, sendo elas o aplicativo e a API. Para o desenvolvimento da API, foi utilizada a linguagem de programação Java, incluindo os frameworks spring boot e lombok. Já a parte do aplicativo, foi desenvolvido na linguagem flutter e para que ele rode no projeto, é necessário executar o seguinte comando no terminal, dentro da pasta onde fica o projeto:

```
flutter pub get
```

A execução do aplicativo em si, é feita no arquivo que possui a extensão “.main”. Já para testar o funcionamento da parte da API, o código deverá ser rodado no arquivo que contém a classe denominada “Application”.

Este projeto poderá ser rodado nas IDE's Visual Studio Code e no IntelliJ IDEA.

Para o armazenamento e manutenção dos dados, foi utilizado um banco de dados relacional PostgreSQL. Contudo, o projeto foi projetado para criar as tabelas de forma automática no momento da execução do programa, sem a necessidade de criá-las no banco usando o comando “Create Table”.

Com isso, as entidades da tabela foram definidas na pasta “entities”, sendo que cada arquivo representa uma tabela e os seus atributos indicam os campos desta entidade. Portanto, no momento da execução da API, a estrutura das tabelas ficará semelhante da forma que foi definida no algoritmo.

Sendo assim, foi utilizado o framework hibernate para a criação e alteração das tabelas no banco de dados, adicionando as seguintes informações no arquivo “application.properties”:

```
spring.jpa.hibernate.ddl-auto: update
spring.jpa.show-sql: true
spring.jpa.generate-ddl: true
```

Além disso, devem ser colocadas neste mesmo arquivo, os dados para se conectar no banco:

```
spring.datasource.url=jdbc:h2:mem:db;DC_CLOSE_DELAY=-1
spring.datasource.username=sa
spring.datasource.password=as
```

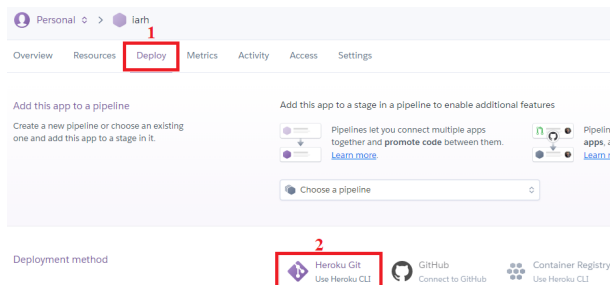
Com todos os pontos mencionados acima estando definidos, qualquer operação feita na API (criar, deletar, alterar, atualizar), será refletida automaticamente no banco de dados. Além disso, é necessário adicionar as seguintes dependências no arquivo "pom.xml", para possibilitar a conexão com o banco PostgreSQL:

```
<dependency>
<groupId>org.postgresql</groupId>
<artifactId>postgresql</artifactId>
</dependency>
```

Os endpoints (requisições http) da aplicação são definidos na pasta controller do projeto da API. Por padrão, o caminho da url é definido como "http://localhost:8080", sendo que a porta 8080, é especificada na variável "server.port" que fica no arquivo de configuração "application-dev.properties".

Para cada classe do Controller, foi adicionada a anotação "@RequestMapping" do spring boot que recebe como parâmetro a url principal da página.

Para subir a API para o heroku, inicialmente foi necessário criar um projeto através do endereço "heroku.com" e, após criado no menu "Deploy" foi selecionada a primeira opção denominada "Heroku Git", sendo necessário instalar o programa Heroku CLI com a finalidade de conectar o git com o heroku.



Logo em seguida, foram executados os seguintes comandos no terminal, no diretório do projeto:

- heroku --version
- heroku login
- heroku git:remote -a nomedoprojeto
- git status
- git add .
- git push heroku master

RESULTADOS OBTIDOS

CONCLUSÃO

Com o trabalho proposto, foi possível aprender como desenvolver um aplicativo móvel, usando boas práticas de programação, além de compreender melhor o funcionamento da inteligência artificial aplicada no processo de seleção de candidatos, principalmente o método da regressão linear.

BIBLIOGRAFIA

- [1] F. Bo-Shone, "Technology, Jobs e the Future of Work in Australia", Julho 2021.
- [2] M. Afonso Paulo, R. Brenno Anderson, A. Cristine Amora e V. Rosângela Couras, "INTELIGÊNCIA ARTIFICIAL - RECURSOS HUMANOS FRENTE AS NOVAS TECNOLOGIAS, POSTURAS E ATRIBUIÇÕES", Outubro 2018.
- [3] T. Ivana, "The Attitude of Job Candidates towards Artificial Intelligence in Hiring Process", Maio 2020.
- [4] S. Abhilasha e S. Apurva, "Impact of Artificial Intelligence on HR practices in the UAE".
- [5] R. Geetha e D. Bhanu Sree, "RECRUITMENT THROUGH ARTIFICIAL INTELLIGENCE: A CONCEPTUAL STUDY".
- [6] J. Mariana Namen, "INTELIGÊNCIA ARTIFICIAL NO RECRUTAMENTO and SELEÇÃO: INOVAÇÃO E SEUS IMPACTOS PARA A GESTÃO DE RECURSOS HUMANOS", Fevereiro 2020.
- [7] F. Krykhtime, A. Morim, N. Vale, L. Fortes e A. Neto, "Aplicando Lógica Fuzzy em um Modelo de Seleção Multicritério para Multi-clientes", Outubro 2013.
- [8] E. Loiola, D. Silveira e M. Bittencourt, "E-Recruiting Utilizando Lógica Fuzzy", Novembro 2022.
- [9] L. Silva, "Modelos de Regressão Fuzzy", 2018.