

Documentação Projeto Integrador III

Módulo 3: Predição e Análise de Desempenho

Clara dos Santos Becker
Erick Augusto Warmling
Lucas Emanuel Gitirana
Lucas Gabriel Falcade Nunes
Marco Antonio Garlini Possamai
Thassyana Snely Alves

SUMÁRIO

1. ESPECIFICAÇÃO DO ESCOPO.....	3
2. DIAGRAMAS DE CASO DE USOS COM CENÁRIO.....	4
3. PROTÓTIPOS NÃO-FUNCIONAIS.....	7
4. DETALHAMENTO DOS RECURSOS.....	13
5. IMPLEMENTAÇÃO DAS TELAS INICIAIS.....	15

1. ESPECIFICAÇÃO DO ESCOPO

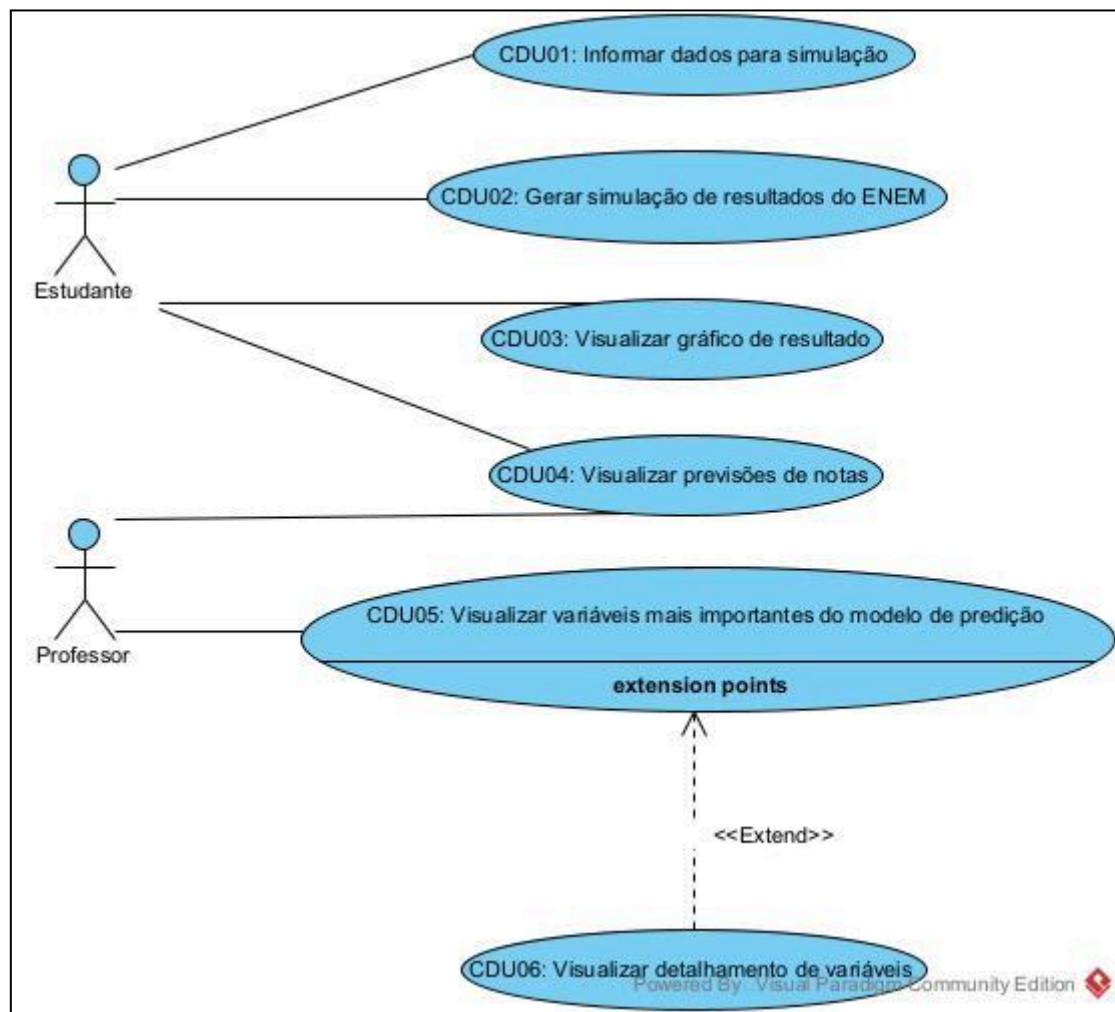
O módulo 3 do sistema de Análise de Microdados do ENEM proposto para a disciplina de Projeto Integrador III tem como objetivo fornecer uma ferramenta de predição do resultado individual ou de uma população com base nos dados socioeconômicos informados pelos participantes das provas no ato de sua inscrição, bem como nas suas notas obtidas. Diante disso, vale ressaltar que a predição será realizada por meio de um modelo preditivo, o qual permitirá identificar relações entre variáveis a fim de determinar quais delas apresentam maior impacto no resultado final.

A partir do processamento realizado pelo modelo, os usuários do sistema poderão interagir com as informações por meio de duas abas: a primeira permitirá a simulação do resultado das notas por meio do preenchimento de um formulário socioeconômico e a segunda permitirá visualizar as variáveis mais influentes do modelo juntamente da justificativa para isso.

No que se refere às ferramentas utilizadas, a tecnologia escolhida em conjunto com as outras equipes para o desenvolvimento do sistema foi o framework Python Streamlit associado ao banco de dados PostgreSQL. O repositório do projeto, por sua vez, pode ser acessado pelo link <https://github.com/lucas-gitirana/analise-microdados-enem-pin-3>.

Ainda, é importante mencionar que durante esta primeira fase do projeto, foram desenvolvidos em conjunto com as outras equipes os métodos de leitura, pré-processamento e armazenamento de dados. No repositório do projeto constam as instruções para a criação do banco de dados, a qual é realizada por meio de um script que lê os arquivos .csv da pasta e realiza o envio para a base de dados. Nas seções a seguir, estão descritos mais detalhes acerca do projeto e da implementação do módulo mencionado.

2. DIAGRAMAS DE CASO DE USOS COM CENÁRIO



CENÁRIOS

CDU 01: Informar dados para simulação

- **Ator:** Estudante
- **Pré-condições:** Estudante autenticado no sistema
- **Fluxo Principal:**
 1. O estudante informa os dados requeridos pelo sistema
 2. O sistema solicita informações
 3. O sistema gera gráficos comparando os dados com as previsões de anos anteriores
- **Fluxo Alternativo:**
Se houver erro de preenchimento, o sistema solicita a correção.
- **Pós-condições:** Os dados do estudante ficam disponíveis para a simulação.

CDU 02: Gerar simulação de resultados do ENEM

- **Ator:** Estudante
- **Pré-condições:** Dados do estudante já informados
- **Fluxo Principal:**

1. O estudante solicita que o sistema gere a simulação
 2. O sistema coleta os dados informados
 3. O modelo de predição é executado
 4. O sistema exibe a nota prevista
- **Fluxo Alternativo:**
Se houver erro de preenchimento, o sistema solicita a correção.
 - **Pós-condição:** A previsão do resultado do estudante é gerada e exibida, o resultado fica disponível para consulta e visualização em gráficos.

CDU 03: Visualizar gráfico de resultado

- **Ator:** Estudante
- **Pré-condições:** O Estudante ter gerado previamente uma simulação de resultado
- **Fluxo Principal:**
 1. O estudante seleciona a opção de visualizar o gráfico.
 2. O sistema gera o gráfico a partir dos dados informados para a simulação.
 3. O gráfico é exibido para o estudante.
- **Fluxo Alternativo:**
Se não houver dados informados, o sistema informa que é necessário gerar uma simulação antes.
- **Pós-condição:** O estudante visualiza o gráfico com os resultados previstos de sua simulação.

CDU 04: Visualizar previsões de notas

- **Ator:** Estudante/Professor
- **Pré-condições:** Existência de uma simulação
- **Fluxo Principal:**
 1. O ator acessa a opção de visualizar previsões de notas.
 2. O sistema recupera os resultados da simulação armazenada.
 3. O sistema exibe as previsões de notas em formato de gráficos.
- **Fluxo Alternativo:**
Se não houver uma simulação disponível, o sistema informa que é necessário gerar a simulação primeiro.
- **Pós-condição:** O ator tem acesso às notas previstas, podendo usá-las para análise ou comparação com metas pessoais.

CDU 05: Visualizar variáveis mais importantes do modelo de predição

- **Ator:** Professor
- **Pré-condições:** Existência de uma simulação.
- **Fluxo Principal:**
 1. O professor solicita a visualização das variáveis importantes.
 2. O sistema processa a relevância das variáveis no modelo preditivo.
 3. O sistema exibe o gráfico com as variáveis ordenadas por importância.
- **Extensão (Extension Point):**
<<extend>> **CDU 06: Visualizar detalhamento de variáveis.**
- **Pós-condição:** O professor visualiza os fatores mais relevantes do modelo.


CDU 06: Visualizar detalhamento de variáveis

- **Ator:** Professor
- **Pré-condições:** Existência de uma simulação.
- **Fluxo Principal:**
 1. O usuário seleciona uma variável importante.
 2. O sistema apresenta o detalhamento da variável selecionada.
- **Fluxo Alternativo:**

O sistema não consegue acessar os dados da variável por algum problema técnico, o sistema exibe uma mensagem de erro.
- **Pós-condição:** O professor visualiza de forma detalhada as variáveis mais importantes.

3. PROTÓTIPOS NÃO-FUNCIONAIS

A seguir, estão elencados os protótipos de tela desenvolvidos para o módulo de predição do sistema. Para conferi-los de forma integral na ferramenta Figma, o seguinte link pode ser acessado: <https://www.figma.com/design/1W2LoEKLHo82kfEnLuYmcl/An%C3%A1lise-de-Microdados-ENEM---PIN-III?node-id=40-631&t=eDDpwWn7Hlp9rEWa-1>.


Predição de Desempenho

Simulação de Resultado
Variáveis Importantes

Preencha os campos socioeconômicos e veja a previsão dinâmica de desempenho em cada área do ENEM.

Resultado da Predição

Linguagens e Códigos

530

Ciências Humanas

702

Ciências da Natureza

670

Matemática

850

Redação

980

Dados do Participante

Sexo

♂ Masculino

♀ Feminino

Prefiro não informar

Idade

0

4

20 anos

19

Renda Familiar

Até 1 SM

1-3 SM

3-5 SM

Mais de 5 SM

Escolaridade do Pai

0

1

4

Ensino Médio

Escolaridade da Mãe

0

4

4

Pós-graduação

Tipo da Escola

Pública

Privada

Federal

Possui acesso à Internet?

Sim

Não

Possui computador?

Sim

Não

Limpar

Gerar Nova Previsão

Visualização Gráfica

Notas Previstas por Área do ENEM

530

702

670

850

980

Linguagens e Códigos

Ciências Humanas

Ciências da Natureza

Matemática

Redação

Área

Linguagens e Códigos

Ciências Humanas

Ciências da Natureza

Matemática

Redação

Nota Prevista

0

200

400

600

Área

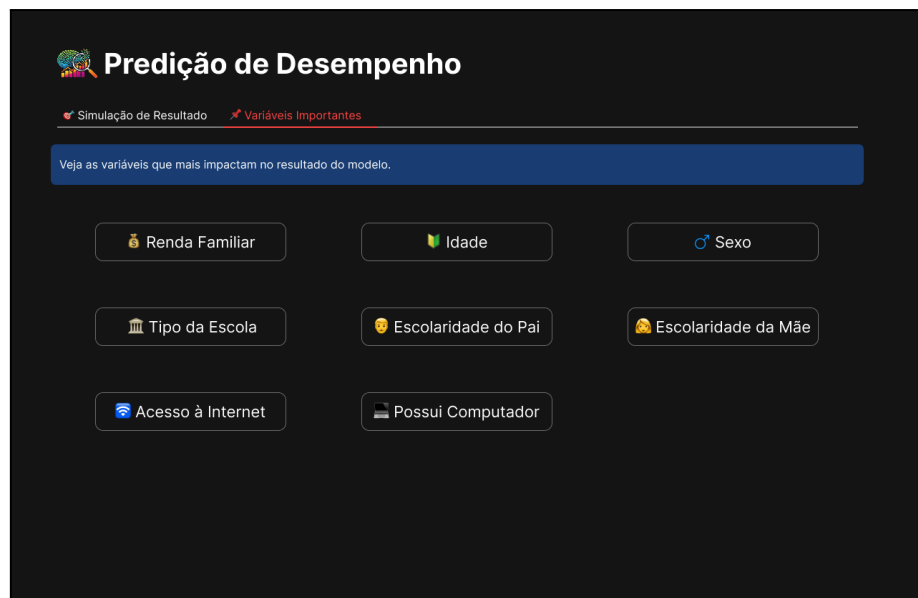
Linguagens e Códigos

Ciências Humanas

Ciências da Natureza

Matemática

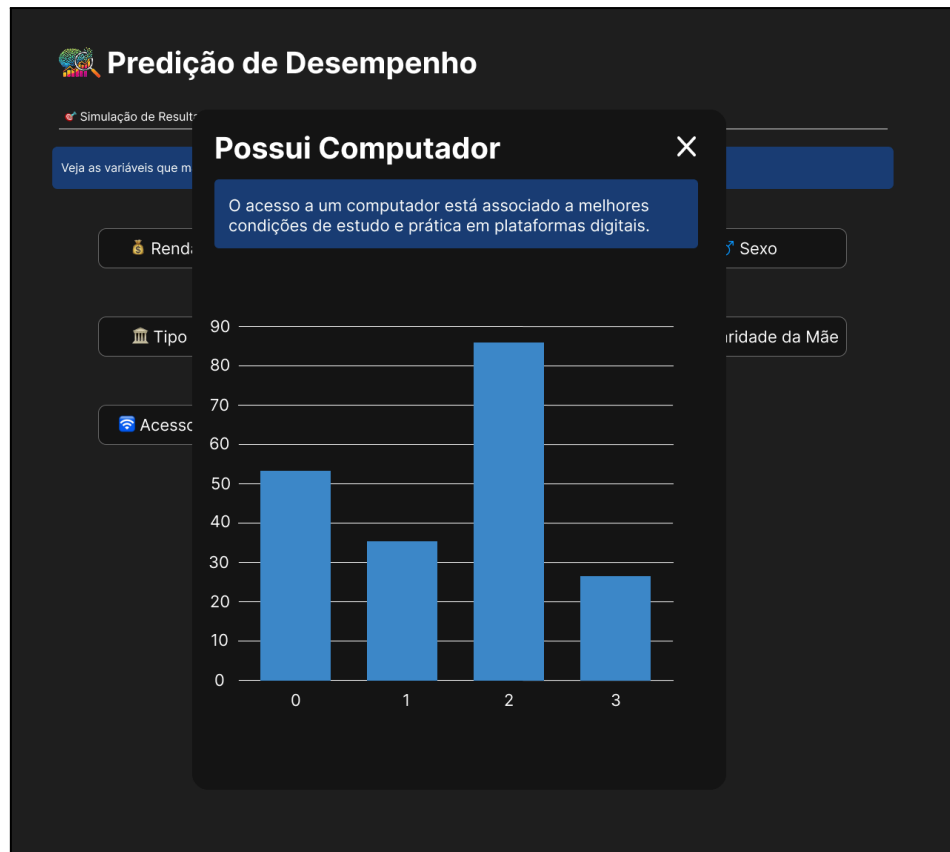
Redação



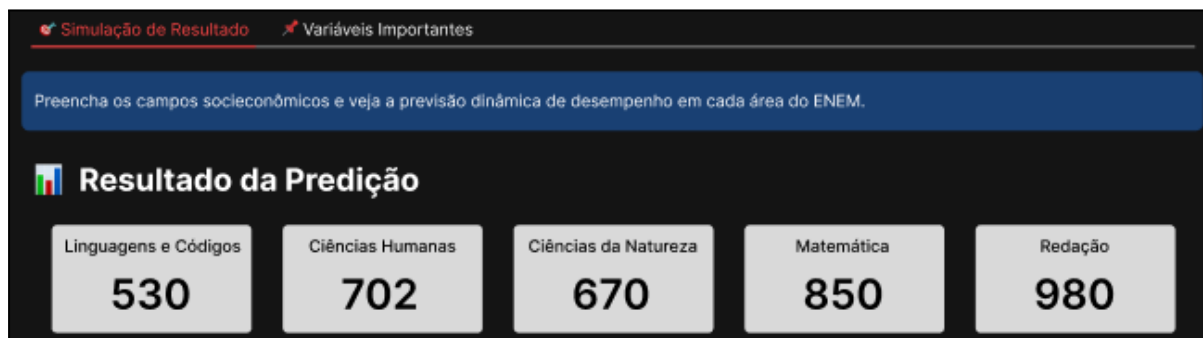




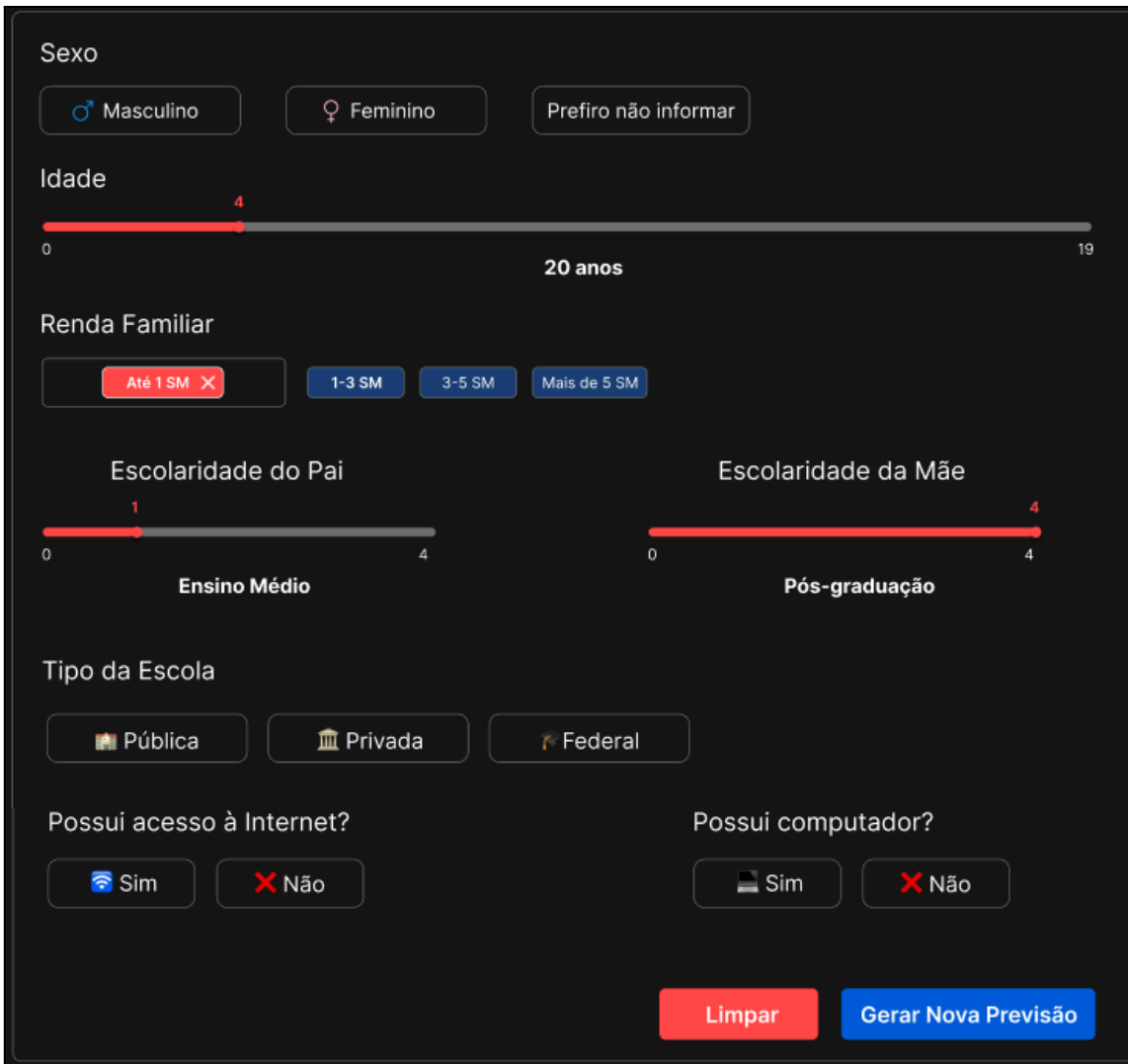




4. DETALHAMENTO DOS RECURSOS



Na parte inicial da tela de predição e desempenho, o sistema irá fornecer duas abas para visualização e interação do usuário. A primeira aba irá solicitar por meio de uma mensagem que a pessoa informe alguns dados respectivos seus em um formulário bem interativo que se encontra mais abaixo dos resultados da predição. A partir desses dados informados, o modelo preditivo treinado pelo nosso grupo irá calcular as possíveis notas que o candidato irá obter em cada área de conhecimento que estrutura a prova do Enem. Inicialmente essa tela irá trazer alguns dados pré-informados, fazendo com que já tenha resultado da predição e a tela não comece vazia.



Sexo

☒ Masculino ☐ Feminino ☐ Prefiro não informar

Idade

0 4 19

20 anos

Renda Familiar

Escolaridade do Pai

0 1 4

Ensino Médio

Escolaridade da Mãe

0 4 4

Pós-graduação

Tipo da Escola

☒ Pública ☐ Privada ☐ Federal

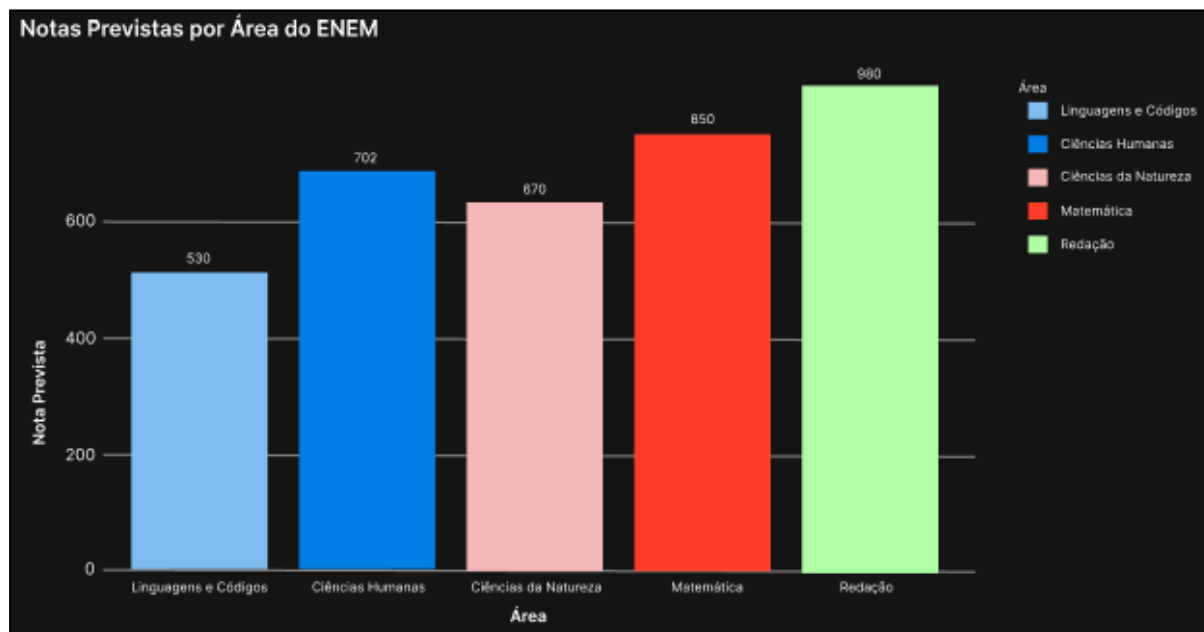
Possui acesso à Internet?

☒ Sim ☐ Não

Possui computador?

☐ Sim ☒ Não

Esse será o formulário interativo que se encontrará logo abaixo do resultado da predição. Nele, o usuário irá informar alguns dos seus dados que até o momento foram considerados mais importantes pelo nosso grupo, onde estes poderão ser atualizados no decorrer do desenvolvimento do sistema e no treinamento do modelo preditivo. Com os dados informados, o usuário poderá então gerar uma nova previsão que fará com que o modelo preditivo recalcule as possíveis notas que o participante poderá obter nas áreas de conhecimento do Enem e atualize a predição no início desta aba.



Por fim, no final dessa primeira aba será possível visualizar as possíveis notas previstas pelo participante em um gráfico bem definido, onde cada área do conhecimento será diferenciada por uma cor e terá a sua definição esclarecida abaixo de cada barra ou ao lado do gráfico.

Simulação de Resultado | **Variáveis Importantes**

Veja as variáveis que mais impactam no resultado do modelo.

Renda Familiar

Idade

Sexo

Tipo da Escola

Escolaridade do Pai

Escolaridade da Mãe

Acesso à Internet

Possui Computador

Na segunda aba referente às variáveis importantes da predição de desempenho, temos os mesmos dados solicitados no formulário da simulação de resultado, onde o

usuário poderá selecionar a variável que desejar e posteriormente será aberto um gráfico demonstrando o nível de importância daquele dado.



Como a imagem demonstra acima, ao clicar em uma variável o sistema irá projetar um gráfico mostrando o nível de significância daquele dado nos estudos e resultados do aluno.

5. IMPLEMENTAÇÃO DAS TELAS INICIAIS

Durante a presente fase do projeto foi desenvolvida de forma coletiva por todos os grupos da disciplina a tela principal do sistema, a qual permite o acesso aos diversos módulos da aplicação. Além disso, foram desenvolvidas também as versões iniciais das telas do módulo 3, que serão aprimoradas conforme o decorrer do projeto. Abaixo, seguem algumas imagens que ilustram os artefatos desenvolvidos.

app

Exploration

Dashboards

Prediction

Questions

gabarito parser

Deploy

Ferramenta de Análise dos Microdados do ENEM

Geral

Utils

Escolha um módulo para começar:

Exploração dos Dados

Dashboards

Predição de Desempenho

Análise de Questões

app

Exploration

Dashboards

Prediction

Questions

gabarito parser

Deploy

Predição de Desempenho

Simulação de Resultado

Variáveis Importantes

Preencha os campos socioeconômicos e veja a previsão dinâmica de desempenho em cada área do ENEM.

Resultado da Predição

Linguagens e Códigos

Ciências Humanas

Ciências da Natureza

Matemática

Redação

502

748

670

506

471

Dados do Participante

Sexo

Masculino

Feminino

Prefiro não informar

Idade

20

0

100

app

Exploration

Dashboards

Prediction

Questions

gabarito parser

Deploy

Resultado da Predição

Linguagens e Códigos

Ciências Humanas

Ciências da Natureza

Matemática

Redação

502

748

670

506

471

Dados do Participante

Sexo

Masculino

Feminino

Prefiro não informar

Idade

20

0

100

Renda Familiar

Até 1 SM

1-3 SM

3-5 SM

Mais de 5 SM

Escolaridade do Pai

Ensino Médio

Fundamental

Não informado

Escolaridade da Mãe

Superior

Fundamental

Não informado

Tipo da Escola

Até 1 SM

1-3 SM

3-5 SM

Mais de 5 SM

16

