

PLANO DE PROJETO

Histórico de Versões

Versão	Data	Autor(es)	Descrição
1.0	09/08/2025	Lucas Garcia de Souza	Elaboração inicial do documento
1.1	15/08/2025	Lucas Garcia de Souza e Luís Augusto	Atualização com base no TAP e cronograma de atividades
1.2	06/10/2025	Lucas Garcia de Souza	Revisão e integração com o Anteprojeto

Autores

- **Lucas Garcia**
- **Luis Augusto**

1. Introdução

Este plano de projeto tem como objetivo detalhar o planejamento, execução e controle do Trabalho de Conclusão de Curso intitulado **“Previsor de Renda no Mercado de Trabalho Brasileiro com Base em Escolaridade e Perfil Socioeconômico”**.

O projeto visa aplicar técnicas de **Inteligência Artificial e Aprendizado de Máquina** sobre bases de microdados públicos, especialmente a **PNAD Contínua**, para estimar de forma confiável a **renda mensal esperada** de trabalhadores brasileiros, considerando fatores como escolaridade, gênero, idade, setor econômico e região geográfica.

2. Justificativa

Apesar da crença generalizada de que a escolaridade determina a renda, dados reais mostram que pessoas com o mesmo nível de instrução apresentam **grandes diferenças salariais**. Isso ocorre devido a múltiplos fatores socioeconômicos e regionais.

O projeto busca, portanto, **compreender e quantificar essa relação**, oferecendo uma ferramenta prática e cientificamente fundamentada para apoiar decisões sobre educação e carreira, além de contribuir para o debate acadêmico sobre **retorno educacional e desigualdade salarial** no Brasil.

3. Referências (Documentos Auxiliares)

- **Rascunho do Anteprojeto de TCC – IFES**, 2025.
 - **TAP – Termo de Abertura do Projeto** – Lucas Garcia & Luís Augusto, 2025.
 - **Gantt Chart – Cronograma de Execução** – Lucas & Luís, 2025.
 - **PNAD Contínua (IBGE)** – Microdados e dicionário técnico, 2023–2024.
 - Kumar (2024), Mullainathan & Spiess (2017), Herrera et al. (2023), Santolin & Oliveira (2023), Pedroso (2023).
-

4. Objetivos do Projeto

Realizar um **estudo aplicado de conclusão de curso** que integre os conhecimentos adquiridos ao longo da graduação em **Sistemas de Informação**, por meio do **desenvolvimento e análise de um modelo computacional baseado em Inteligência Artificial**. O projeto tem como finalidade **comprovar a capacidade técnica, científica e gerencial dos discentes**, cumprindo os requisitos necessários para a **obtenção do grau de Bacharel em Sistemas de Informação** pelo Instituto Federal do Espírito Santo – Campus Colatina.

5. Escopo

O projeto entregará:

- Previsor de IA treinado para estimar a empregabilidade com base no perfil do usuário.
 - Relatório técnico (Artigo) com metodologia, análise e resultados.
 - Documentação de código.
 - Apresentação em slides.
-

6. Planos de Entregas do Projeto

O plano de entregas define todas as etapas, produtos e prazos do projeto, englobando o desenvolvimento técnico, a produção documental e a apresentação final. Cada entrega deve atender aos critérios de qualidade, funcionalidade e conformidade acadêmica.

As principais entregas são:

- 1. Planejamento e Organização (Outubro/2025):**
Elaboração e aprovação, pelo orientador, do documento de escopo, justificativas, cronograma e Termo de Abertura do Projeto (TAP).
- 2. Revisão Bibliográfica (Dezembro/2025):**
Produção da fundamentação teórica consolidada, com referências atualizadas e devidamente normalizadas segundo as normas da ABNT.
- 3. Coleta e Preparação de Dados (Março/2026):**
Tratamento, validação e documentação da base de dados PNAD.
- 4. Modelagem Preditiva (Junho/2026):**
Desenvolvimento e comparação de modelos de Inteligência Artificial (Regressão, Random Forest e XGBoost), com desempenho mínimo de $R^2 \geq 0,70$ e $MAE \leq 500$.
- 5. Backend/API (Julho/2026):**
Implementação de um serviço de previsão funcional, com rotas, logs e testes devidamente documentados.

6. Frontend/Dashboard (Setembro/2026):

Desenvolvimento de interface em React.js contendo formulários, gráficos explicativos e visualização de resultados.

7. Integração e Testes (Outubro/2026):

Validação do sistema completo, com testes de desempenho e usabilidade.

8. Documentação do Código (Novembro/2026):

Entrega do repositório no GitHub, contendo instruções de execução, arquitetura do sistema e comentários de código.

9. Artigo Técnico (Novembro/2026):

Produção de relatório científico formatado conforme as normas da ABNT, apresentando metodologia, resultados obtidos e conclusões.

10. Apresentação e Defesa (Dezembro/2026):

Elaboração dos slides de defesa, demonstração funcional do sistema e exposição oral perante a banca avaliadora.

7. Exclusões Específicas (Não Escopo)

O projeto não compreende:

- Implantação do sistema em produção para uso contínuo por terceiros.
 - Integração direta com sistemas externos ou bancos de dados corporativos.
 - Manutenção contínua da aplicação após a entrega do TCC.
-

8. Premissas

- As bases de dados públicas utilizadas (PNAD Contínua, RAIS, Censo Demográfico, CAGED) continuarão sendo coletadas e disponibilizadas por obrigação legal e política pública, conforme legislação vigente.
 - Os microdados permanecerão de acesso livre e gratuito durante todo o período de execução do projeto.
 - A infraestrutura mínima necessária (computador pessoal e acesso à internet) estará disponível durante o desenvolvimento.
-

9. Restrições

- O projeto é limitado ao uso de dados públicos disponíveis, sem aquisição de bases privadas ou pagas.
 - O prazo de entrega está condicionado ao calendário acadêmico do TCC.
 - Equipe restrita aos membros do projeto, não podendo haver mais contratações ou custos com recursos humanos.
 - Competência técnica dos membros da equipe, não demandando conhecimentos adicionais com o intuito de garantir a entrega sem adiamento da entrega.
-

10. Tecnologias Utilizadas

Camada	Tecnologia
Coleta e Processamento	Python 3.10, Pandas, NumPy, BigQuery, Jupyter Notebook
Modelagem IA	scikit-learn, XGBoost, SHAP
Backend/API	Flask, FastAPI, PostgreSQL
Frontend	React.js, Vite, TailwindCSS
Visualização de Dados	Matplotlib, Plotly, Recharts
Controle de Versão	Git/GitHub
Infraestrutura e Deploy (demo)	Render / Railway / AWS Sandbox

11. Cronograma de Entregas (Síntese)

Com base no Gantt Chart.

Fase	Período	Entregas Principais

Planejamento e Organização	Ago–Out/2025	Declaração de Escopo, Requisitos e Revisão Bibliográfica
Coleta e Preparação de Dados	Fev–Mar/2026	Dataset tratado, normalizado e validado
Modelagem e Treinamento	Abr–Jun/2026	Modelos IA treinados e avaliados
Backend/API	Jun–Jul/2026	API funcional com documentação
Frontend e Dashboard	Jul–Set/2026	Interface com gráficos e previsor ativo
Integração e Testes	Out–Nov/2026	Testes de desempenho e usabilidade
Documentação e Defesa	Dez/2026	Relatório final e apresentação pública

Para mais detalhes segue o documento separado.

12. Descrição de Tarefas (Principais)

1. **Revisão bibliográfica e fundamentação teórica**
2. **Tratamento dos microdados da PNAD**
3. **Seleção e engenharia de variáveis**
4. **Treinamento e validação de modelos de IA**
5. **Desenvolvimento da API preditiva**
6. **Criação do frontend interativo e dashboard**
7. **Testes de integração e desempenho**
8. **Redação do relatório e preparação da defesa**

Para mais detalhes segue o documento separado.

13. Levantamento de Requisitos

Em documento de Anexo separado

14. Descrição de Funcionalidades

- 1. Previsor de Renda:** formulário com variáveis de entrada.
 - 2. Visualização Explicativa:** mostra quais variáveis influenciam a renda.
 - 3. Exportação de Resultados:** permite baixar o relatório personalizado.
-

15. Orçamento (Custo Estimado)

Baseado no TAP.

Tipo de Custo	Descrição	Valor Estimado
Hospedagem temporária	Deploy de demonstração (Render/AWS)	R\$ 200,00
Trabalho técnico (1.500h × R\$25/h)	Desenvolvimento e pesquisa	R\$ 37.500,00
Total Estimado		R\$ 37.700,00

16. Riscos do Projeto

Identificação dos riscos: realizada com base na experiência dos desenvolvedores, no histórico de TCCs anteriores e nas características técnicas do projeto.

Análise qualitativa: classifica cada risco conforme probabilidade (Baixa, Média, Alta) e impacto (Baixo, Médio, Alto).

Planejamento de resposta: define ações preventivas e corretivas para cada risco identificado.

Monitoramento e controle: os riscos serão revisados quinzenalmente nas reuniões com o orientador, e o plano será ajustado conforme a evolução do cronograma.

ID	Risco	Causa Provável	Conseqüência Potencial	Probabilidade	Impacto	Indicadores de Ocorrência	Estratégia / Ação Preventiva	Ação Corretiva	Responsável
R1	Atraso devido ao cronograma acadêmico	Acúmulo de tarefas e sobreposição de prazos em disciplinas	Atraso nas entregas e comprometimento da qualidade do produto	Alta	Alto	Atividades em atraso no Gantt ou falhas em checkpoints	Reuniões semanais de acompanhamento e controle de prazos	Replanejar atividades e redistribuir tarefas	Lucas G. / Pedro E.
R2	Dados incompletos ou inconsistentes	Falhas na base PNAD, variáveis nulas ou formatação incorreta	Resultados incorretos ou impossibilidade de treinar o modelo	Média	Médio	Alto volume de NaN ou inconsistências detectadas no pré-processamento	Aplicar técnicas de <i>data cleaning</i> e normalização	Substituir variáveis problemáticas ou buscar fontes secundárias (RAIS, CAGED)	Lucas G.
R3	Desistência de membro	Sobrecarga ou falta de disponibilidade de tempo	Retrabalho e atraso em atividades-chave	Baixa	Alto	Falta em reuniões e não cumprimento de prazos	Divisão redundante de tarefas críticas e documentação detalhada	Redistribuir funções e revisar cronograma	Orientador / Lucas G.
R4	Dificuldades técnicas com	Falta de experiência ou falhas de	Interrupção no desenvolvimento	Média	Médio	Erros de execução	Estudo prévio e testes incrementais	Solicitar apoio técnico e	Pedro E. / Lucas G.

	bibliotecas de IA	compatibilidade de versão	do modelo preditivo			frequentes ou baixo desempenho	ais das bibliotecas	substituição por métodos equivalentes	
R5	Perda de dados do projeto	Falha de hardware, erro humano ou repositório corrompido	Perda de código, datasets e resultados	Baixa	Alto	Arquivos inacessíveis ou corrompidos	Backups automáticos no GitHub e Google Drive	Recuperar versões anteriores via Git e backups locais	Pedro E. / Lucas G.
R6	Baixa acurácia do modelo de IA	Escolha inadequada de algoritmo ou overfitting	Resultados inconsistentes e baixa confiabilidade	Média	Alto	$R^2 < 0,7$ e MAE > 500 nas métricas	Testar múltiplos algoritmos e ajustar hiperparâmetros	Reavaliar features e realizar <i>tuning</i> adicional	Pedro E.
R7	Problemas de integração entre backend e frontend	Diferença de formatos entre API e interface	Sistema instável ou resultados incorretos na visualização	Média	Médio	Falhas na comunicação API (status 500/404)	Padronização de contratos e testes de integração	Revisar rotas e formato de resposta da API	Lucas G.
R8	Sobrecarga de hardware local	Limitações de processamento durante o treinamento	Lentidão ou travamentos durante execução	Média	Baixo	Execuções longas ou crashes em notebooks	Uso de <i>cloud computing</i> (Google Colab, AWS Sandbox)	Reducir batch size ou usar amostra menor	Pedro E.

R9	Falhas na apresentação final	Problemas técnicos ou perda de material	Comprometimento da defesa pública	Baixa	Alto	Falhas em simulações prévias	Testes de apresentação e cópia de segurança em pen drive e nuvem	Substituir equipamento e garantir backup completo	Lucas G.
----	------------------------------	---	-----------------------------------	-------	------	------------------------------	--	---	----------

Frequência: Revisão quinzenal dos riscos em reunião com o orientador.

Ferramentas: Planilha de controle (Google Sheets) integrada ao cronograma Gantt.

Indicadores-chave: porcentagem de riscos ativos, número de riscos mitigados, e impacto acumulado.

Critério de encerramento: risco será considerado encerrado quando a probabilidade e o impacto forem rebaixados simultaneamente para “Baixo”.

17. Gerenciamento dos Riscos

Objetivo

Assegurar que todos os riscos identificados sejam acompanhados e controlados ao longo do ciclo de vida do projeto, minimizando seus impactos e maximizando as oportunidades de sucesso das entregas.

Responsabilidades

- **Gerente do Projeto (Lucas Garcia de Souza):** responsável por supervisionar todo o plano de riscos, atualizar os status e acionar as ações corretivas necessárias.
- **Equipe de Desenvolvimento:** responsável por identificar novos riscos, relatar ocorrências e aplicar medidas preventivas.

- **Orientador:** atua como agente de controle e validação das decisões técnicas e de mitigação.
- **Facilitador / Professor da Disciplina:** supervisiona o alinhamento metodológico do gerenciamento dos riscos ao PMBOK.

Procedimentos de Monitoramento

O monitoramento dos riscos será realizado conforme as etapas a seguir:

1. Acompanhamento Contínuo:

- Utilização de uma **planilha de controle de riscos (Risk Register)** hospedada em Google Sheets e sincronizada com o cronograma Gantt.
- Atualizações semanais ou sempre que houver alteração de probabilidade, impacto ou status do risco.

2. Reuniões de Avaliação Quinzenais:

- Reuniões com a equipe e o orientador para revisão dos riscos ativos, verificação de novos riscos e avaliação das respostas aplicadas.
- Cada reunião incluirá uma análise de desempenho do plano de mitigação e definição de ações preventivas adicionais.

3. Revisões de Progresso Alinhadas ao Gantt Chart:

- Acompanhamento dos marcos (milestones) do projeto para identificar atrasos, gargalos ou desvios que indiquem materialização de riscos.
- As revisões serão registradas no histórico do projeto e servirão de base para ajustes de cronograma.

4. Indicadores de Monitoramento:

- Percentual de riscos mitigados versus ativos.
- Número de riscos novos identificados por mês.
- Grau médio de exposição do projeto (média ponderada de impacto × probabilidade).

Procedimentos de Resposta

Em caso de risco identificado ou materializado, serão adotadas as seguintes medidas:

- **Ação Preventiva:** executada imediatamente ao detectar sinais de alerta (indicadores de risco), buscando evitar a ocorrência do evento.
- **Ação Corretiva:** aplicada após a materialização do risco, com foco em minimizar o impacto e restabelecer o cronograma planejado.
- **Escalonamento:** riscos classificados como *Alta Probabilidade + Alto Impacto* serão comunicados ao orientador e reavaliados em até 48 horas.
- **Replanejamento Imediato:** se o risco afetar o caminho crítico do projeto, será feita uma reprogramação parcial no Gantt Chart, priorizando a retomada das atividades críticas.

Ferramentas e Documentos Utilizados

- **Planilha de Controle de Riscos (Risk Log)** – acompanhamento contínuo de status, causas e medidas.
- **Gantt Chart** – monitoramento visual do cronograma e correlação com eventos de risco.
- **Relatórios de Progresso Quinzenais** – consolidação de decisões e histórico de mitigação.
- **GitHub Issues** – registro técnico de falhas, bugs e eventos de risco operacional.

Comunicação e Atualização

- Os riscos e ações associadas serão comunicados via **reuniões quinzenais, WhatsApp** (para riscos urgentes) e **GitHub/Notion** (para documentação oficial).
 - Toda atualização relevante será registrada no **histórico do plano de projeto** e revisada em conjunto com o orientador.
 - Novos riscos identificados durante o desenvolvimento serão adicionados ao registro, com atribuição imediata de responsável e estratégia de mitigação.
-

18. Plano e Gerenciamento de Comunicação

Canal	Frequência	Responsável	Objetivo
Reuniões presenciais	Quinzenal	Equipe e orientador	Acompanhamento técnico
E-mail / WhatsApp	Semanal	Equipe	Comunicação rápida e decisões operacionais
GitHub / Notion	Contínuo	Equipe	Controle de versão e documentação
Relatórios parciais	Mensal	Equipe	Progresso e atualização do orientador

19. Critérios de Aceitação / Qualidade

Os critérios de aceitação e qualidade foram definidos para garantir que o produto final atenda às expectativas funcionais, técnicas e acadêmicas do projeto. Esses critérios servirão como base para a validação do software e para a avaliação geral do TCC.

1. Qualidade Técnica (Modelo de Inteligência Artificial)

- **Precisão ($R^2 \geq 0,70$):** o modelo de IA deve apresentar coeficiente de determinação igual ou superior a 0,70 nas previsões de renda.
- **Erro Médio Absoluto ($MAE \leq 500$):** o erro médio entre o valor previsto e o real não deve ultrapassar R\$ 500,00.
- **Erro Quadrático Médio ($RMSE \leq 800$):** garante estabilidade das previsões em faixas médias de renda.
- **Explicabilidade (SHAP):** cada previsão deve ser acompanhada por um gráfico explicativo de importância das variáveis (SHAP values).
- **Validação Cruzada:** deve haver evidência de que o modelo foi validado em múltiplos conjuntos de dados (cross-validation ≥ 5 folds).

2. Qualidade de Software

- **Performance:** tempo de resposta máximo de 3 segundos para requisições de previsão.
- **Disponibilidade:** O sistema deve permanecer funcional durante os testes de defesa e demonstração.
- **Usabilidade:** a interface deve ser responsiva, intuitiva e compatível com navegadores modernos e dispositivos móveis.
- **Segurança:** não deve haver exposição de dados pessoais ou sensíveis (uso exclusivo de dados anonimizados).
- **Confiabilidade:** todos os cálculos e previsões devem ser reproduzíveis com a mesma entrada.

3. Qualidade de Documentação

- **Clareza e completude:** o relatório técnico deve conter metodologia, justificativas, testes e resultados de forma comprehensível.
- **Reprodutibilidade:** o código e instruções devem permitir replicar os experimentos em ambiente local ou em nuvem.
- **Formatação acadêmica:** o texto deve seguir as normas ABNT e o formato exigido pelo IFES.

4. Critérios de Aceitação Final

O projeto será considerado **aceito e concluído com sucesso** se:

1. O protótipo funcional realiza previsões corretas com base nos critérios acima.
2. O modelo apresentou resultados estatisticamente satisfatórios ($R^2 \geq 0,70$ e $MAE \leq 500$).
3. As funcionalidades previstas estão totalmente implementadas.
4. O relatório técnico foi entregue dentro do prazo, validado pelo orientador e pela banca.

5. O sistema demonstra valor prático e coerência com os objetivos propostos no anteprojeto.