De: Marina Ferrari

Para: A equipe de caso da Nature Conservancy

Assunto: [Novo Projeto] Projeto de mitigação do impacto climático

Olá, equipe, espero que esteja tudo bem.

A diretora de estratégia da The Nature Conservancy (TNC) entrou em contato conosco porque está muito impressionada com nossos materiais de práticas sobre mudanças climáticas e projetos. Ela solicitou nossa assistência para projetar um conjunto de iniciativas estratégicas para mitigar o impacto climático da agricultura brasileira e pecuária na região amazônica (especificamente no estado do Pará), mantendo o crescimento desse setor.

Eu dividi todo o projeto em fluxos de trabalho, seguindo a estrutura da Bain para elaborar um Plano Estratégico. Você encontrará no slide anexo a estrutura contendo as principais perguntas que queremos responder para cada fluxo de trabalho.

Também consolidei abaixo as metas específicas e os insights esperados para cada um dos nossos fluxos de trabalho para que você tenha uma noção de como vamos cumprir o objetivo geral do projeto. Vou atribuir uma análise específica para você em um dos fluxos de trabalho assim que eu validar o plano de trabalho com nosso Parceiro Bain.

• **Workstream 1 (Case for Change & Ambition):** Gostaríamos de entender e quantificar o impacto que a agricultura e a pecuária tiveram na região amazônica e calcular o potencial do nosso projeto de redução de emissões de carbono equivalente no estado do Pará.

• **Workstream 2 (Where to Play):** O objetivo deste fluxo de trabalho é definir quais regiões e tipos de produtos devem ser o foco da TNC – o ideal é que sejam áreas com alta demanda do setor de agronegócios e com deterioração das condições ambientais.

• **Workstream 3 (How to Win):** Neste workstream devemos mapear as alavancas mais importantes para mitigar o impacto ambiental do agronegócio, ajudando a TNC a cumprir a ambição do projeto.

• **Workstream 4 (Implementation roadmap & Enablers):** Com base nos insights anteriores, este fluxo de trabalho criará o conjunto de estratégias iniciativas que a TNC realizará para executar o plano e acompanhar seus resultados. Também mapeamos os facilitadores necessários para garantir um sucesso implementação do plano.

Obrigada,

Marina.

**Estrutura do Plano Estratégico e perguntas-chave**

**Case for Change**

• Por que precisamos mitigar o impacto ambiental da agricultura e pecuária na região amazônica do Brasil?

• Qual enunciado do problema estamos tentando resolver?

**Ambition**

• Qual é a ambição do plano no Pará, em termos de redução de emissões de carbono equivalente? – Como atender a demanda prevista para agricultura e pecuária, minimizando as emissões de carbono equivalentes?

• Em que prazo o projeto deve capturar essa ambição?

**Where to Play**

• Quais combinações de região + produto do agronegócio deve ser priorizado? Estas são áreas prioritárias para investir:

– Quais regiões têm maior probabilidade de aumentar a área dedicada à agricultura e pecuária?

– Quais regiões têm a maior taxa de desmatamento?

– Quais regiões têm a menor produtividade?

– Quais produtos (tipos de culturas e gado) têm a maior pegada ambiental?

**How to Win**

• Quais são as formas de mitigar o impacto ambiental da agricultura e pecuária no Brasil?

**Implementation roadmap**

• Qual é o roteiro de implementação do plano estratégico?

• Que conjunto de iniciativas são globais (aplicadas a todas as regiões e produtos) versus específicas (aplicadas a uma região, produto e/ou tamanho de propriedade rural específicos)?

• Como cada iniciativa contribui para alcançar a ambição do projeto?

• Como monitorar a eficácia do plano proposto? Quais são os indicadores de sucesso (KPIs) do projeto?

**Enablers**

• Quais são os facilitadores mais relevantes para garantir uma implementação bem-sucedida das iniciativas estratégicas?

**Você recebeu sua primeira atribuição de projeto do seu gerente**

De: Marina Ferrari

Para: A equipe de caso da Nature Conservancy

Assunto: Atribuição da Fase 1

Olá, novamente equipe!

Espero que você tenha tido a chance de ler o e-mail de apresentação do caso para ter uma visão geral do nosso novo projeto. Gostaria de solicitar sua ajuda com o fluxo de trabalho "**Where to Play**". Como mencionei, o objetivo deste fluxo de trabalho é avaliar quais são as áreas prioritárias (em termos de região e tipo de produto) para investir os recursos da TNC. Essas áreas devem ser as mais afetadas pela expansão do agronegócio, em termos de impacto ambiental.

Para realizar essa análise, planejamos investigar 4 critérios diferentes:

**1. Área de terra prevista:** nossa hipótese é que as regiões com maior demanda esperada para o agronegócio devem ser priorizadas para garantir que a demanda seja atendida, minimizando o impacto ambiental;

**2. Taxa de desmatamento:** áreas com maior taxa de desmatamento são mais propensas a prejudicar o meio ambiente e devem ser protegidas;

**3. Produtividade:** regiões com menor produtividade demandam mais área para entregar a mesma quantidade de produto, o que poderia definir mais pressão para abrir novas áreas sobre florestas nativas;

**4. Classificação de produtos (pegada ambiental):** alguns produtos (tipos de culturas e pecuária) consomem mais recursos naturais e emitem mais gases de efeito estufa do que outros – portanto, aqueles com maior pegada ambiental também devem ser priorizados em nosso plano.

Peço sua ajuda no 1º critério: seu objetivo é criar um modelo que preveja a área de terra destinada ao plantio de culturas agrícolas e a pastagens para pecuária, no estado do Pará. Com esta importante contribuição, teremos uma melhor noção de onde focar a seguir.

Já alinhei este tópico com Research & Data Services (RDS) e eles conseguiram reunir alguns dados que podem nos ajudar. Danilo do O Advanced Analytics Group (AAG) já trabalhou com esses bancos de dados antes, fale com ele para que ele possa explicar um pouco melhor sobre as bases de dados e se você pode aproveitar alguma análise que ele já tenha feito.

O tempo é curto, então recomendo fortemente que você divida o trabalho.

Muito obrigado, equipe e boa sorte!

Obrigada,

Marina.

**Um colega do Advanced Analytics Group tem experiência com os bancos de dados e pode apoiá-lo**

De: Danilo Carvalho

Para: Equipe do caso The Nature Conservancy - modelo de previsão de fluxo de trabalho

Assunto: [Novo Projeto] Bancos de dados de agricultura

Olá, pessoal,

Temos um arquivo de banco de dados chamado historic-database.csv. Ele contém as séries temporais com valores conhecidos até 2017. A tabela de dados é a seguinte:

Tabela

Descrição gerada automaticamente

Coluna **year**: o ano desta linha de dados.

Coluna **city\_code**: o código da cidade desta linha de dados (dentro do estado do Pará).

Coluna **product\_type**: tipo do produto do agronegócio desta linha (lavouras permanentes, culturas temporárias ou pecuária).

Coluna **product**: o produto específico desta linha.

Coluna **destinated\_area**: a área destinada ao cultivo ou produção deste produto, neste município, este ano, em hectares.

Seu modelo deve ser capaz de prever a área destinada para cada produto e cidade, para os anos de 2018 e 2019. Por favor, também escreva 1-2 parágrafos com os principais insights e sua interpretação dos resultados do modelo.

Também anexei uma documentação técnica para apoiá-lo nesta análise. Ele explica como você pode enviar seus resultados para a AAG e como nós vamos avaliar seu modelo.

Obrigado,

Danilo.

**Existe um conjunto de diretrizes para a análise da Fase 1 (1/3)**

• Grupos de participantes receberão os bancos de dados e poderão trabalhar na construção de seus modelos a partir de então. Sugerimos que os candidatos usam os anos de 2016 e 2017 como dados de teste para criar seus modelos.

• Periodicamente, os candidatos poderão submeter os resultados obtidos nas suas bases de provas (anos 2016 e 2017). A organização do O Desafio Estratégico calculará o desempenho de cada grupo e fornecerá um ranking atualizado da competição. Esta classificação refere-se apenas para testar dados e tem como objetivo colocar os grupos de participantes em termos de desempenho em relação a outros concorrentes.

• Ao final do período da primeira fase, cada grupo deverá apresentar os valores previstos para os anos de 2018 e 2019. Esses valores serão utilizados para calcular o desempenho final. Apenas o resultado do desempenho final, calculado sobre as previsões para os anos de 2018 e 2019, será utilizado para a classificação final – um dos critérios considerados para definir os aprovados para a segunda fase.

**Existe um conjunto de diretrizes para a análise da Fase 1 (2/3)**

**Formato de submissão final Diretrizes adicionais**

• A submissão dos resultados deve ser feita em arquivo csv, contendo as mesmas colunas como o arquivo historic-database.csv. O conteúdo da coluna **destinated\_area** deve ser os valores calculados pelo modelo de previsão desenvolvido pelos participantes para os anos de referência (2016 e 2017 no teste e 2018 e 2019 na submissão final).

Tabela

Descrição gerada automaticamente

• Além disso, a equipe deve enviar o(s) arquivo(s) que contém todo o código utilizado para criar o modelo. Os arquivos de código podem ser scripts ou notebooks em Python ou R idiomas, ou mesmo arquivos Excel onde o modelo foi desenvolvido.

• Inclua também um breve resumo (1-2 parágrafos) com seus insights e interpretação dos resultados do modelo (formato word ou ppt).

**Diretrizes adicionais**

• Para garantir a reprodutibilidade, sugerimos que as sementes aleatórias sejam definidas manualmente. Assim, executar o código em um computador diferente daquele que gerou o modelo não apresentará resultados diferentes. Recomendamos que os participantes testem seus códigos em mais de um computador para garantir que os resultados sejam idênticos antes da apresentação final.

• Assegurar a compreensão e avaliação do código, bem como a lógica por trás do desenvolvimento do modelo, sugerimos que os participantes insiram comentários ao longo do código, descrevendo o que está sendo implementado em cada seção, qual é o objetivo desta implementação, quaisquer instalações utilizadas e, principalmente, quais os motivos das escolhas dos parâmetros.

• Também recomendamos que as etapas de análise de dados antes da criação do modelo (visualização, filtros, análise de distribuição, etc.) resultados sejam comentados. Desta forma, a equipa do Desafio Estratégico pode de facto garantir que os resultados foram obtidos através da aplicação correta de técnicas de ciência de dados.

**Existe um conjunto de diretrizes para a análise da Fase 1 (3/3)**

Apenas 10 grupos com melhor desempenho serão selecionados para a próxima fase. Existem dois critérios principais de avaliação:

1. Resultados médios do WMAPE

– Grupos com melhores resultados médios WMAPE de cultivos permanentes, cultivos temporários e pastagens terão seus modelos avaliados e reproduzidos pela equipe do Desafio Estratégico. Aqueles que apresentam o mesmo resultado apresentado na execução do código pela equipe do Desafio Estratégico e que tiver justificado todas as escolhas de parâmetros receberá maior pontuação neste critério.

– Portanto, é importante destacar que todos os resultados devem ser reprodutíveis e compreensíveis pela Equipe Desafio Estratégico.

– A métrica de desempenho será WMAPE (Erro de Porcentagem Absoluta Média Ponderada). Um resultado WMAPE será calculado para cada tipo de produto (temporário, permanente e pastagem) e o resultado final. Resultado do grupo será a média desses três valores WMAPE. O WMAPE é calculado pela seguinte fórmula:

Texto

Descrição gerada automaticamente

– Onde A é o vetor que contém os valores reais conhecidos e F é o vetor que contém as previsões feitas pelos participantes.

– Ressaltamos que o numerador da fórmula se refere à soma dos módulos das diferenças entre o valor real e o valor previsto, portanto erros de mesma magnitude, mas com sinais inversos, não se cancelam.

2. Abordagem utilizada para desenvolver o modelo

– Os grupos com resultados médios fortes do WMAPE também terão seus modelos avaliados em termos de nível de complexidade, uso de técnicas inovadoras de ciência de dados e qualidade dos comentários e dados etapas de análise/avaliação utilizadas para desenvolver o modelo. As equipes que demonstrarem ideias inovadoras e/ou dedicação significativa para desenvolver a solução receberão pontuação mais alta neste critério.