

CAÇADA INTELIGENTE – UM JOGO PARA ESTIMULAR O APRENDIZADO SOBRE UMA PLATAFORMA DE IA

SMART HUNT – A GAME TO STIMULATE LEARNING ABOUT AN AI PLATFORM

Graziella Bonadia¹, Fabiani Souza², Gilmar Barreto³

DOI: 10.37702/REE2236-0158.v41p166-176.2022

RESUMO

Este trabalho apresenta o processo de criação de um jogo cujo objetivo foi o de acelerar o aprendizado sobre o uso da plataforma de Inteligência Artificial (IA) desenvolvida pelo CPQD e de aumentar o interesse de cientistas de dados e analistas pela tecnologia. Além disso, esta abordagem foi colocada em teste para que, posteriormente, pudesse ser adaptada e disponibilizada de maneira *on-line* e com respostas automáticas. O jogo Caçada Inteligente foi realizado com um público com interesse no tema IA, mas que, contudo, não era de especialistas na área. Eles puderam acessar a plataforma para cumprir as tarefas solicitadas. De maneira geral, a avaliação do jogo mostrou que ele foi considerado estimulante e uma boa forma para o aprendizado da nova tecnologia.

Palavras-chave: gamificação; plataforma; Inteligência Artificial; Machine Learning.

ABSTRACT

This paper presents the creation process of a game that seeks to accelerate the learning curve of the use of an AI platform developed by CPQD and to raise the attention of data scientists and analysts to the technology. In addition, this approach was tested so that it can later be adapted and made available online with automatic replies. The game, Smart Hunt, was executed enthusiasts but not AI experts. They were able to access the platform to fulfill the requested tasks. Overall, an evaluation of the game by some participants showed that it was stimulating and a good way to learn the new technology.

Keywords: gamification; platform; Artificial Intelligence; Machine Learning.

INTRODUÇÃO

Em 2019, iniciou-se o desenvolvimento de uma Plataforma de Inteligência Artificial (IA) voltada para a aceleração da construção de modelos de *machine learning* e sua implantação por pessoas que não são, necessariamente, da área de Ciência de Dados. Esse desenvolvimento privilegiou questões envolvendo usabilidade para que a curva de

aprendizado fosse mais rápida ao ser comparada com outras tecnologias.

De forma geral, realizar os primeiros passos em uma nova tecnologia sempre é um desafio. Além disso, é natural que em muitos casos as novas tecnologias não consigam obter adesão massiva por apresentarem uma forte barreira quanto à dificuldade de entendimento dos primeiros passos.

Com isso, além dos aspectos de usabilidade, o projeto previu uma série de

¹ Doutoranda em Engenharia Elétrica pela Universidade Estadual de Campinas (UNICAMP), pesquisadora do CPQD; bonadia@cpqd.com.br

² Pesquisadora do CPQD; fabiani@cpqd.com.br

³ Prof. Dr. da Universidade Estadual de Campinas; gbarreto@dsif.fee.unicamp.br

tutoriais com vídeos curtos e mensagens diretas, adaptando-se ao formato que, atualmente, é mais aceito pelas pessoas. Ainda assim, existe uma barreira quanto a fazer com que as pessoas identifiquem essa nova tecnologia como interessante para incluir na sua mala de ferramentas. Dessa forma, resta a questão de como tornar instigante o aprendizado sobre uma nova tecnologia que pode estar fora da zona de conforto dos usuários potenciais.

A plataforma de IA em questão foi desenvolvida pelo CPQD⁴, com apoio do Ministério da Ciência, Tecnologia e Inovação (MCTI) e da Rede Nacional de Ensino e Pesquisa (RNP), e tem como público-alvo dois perfis principais: (i) pessoas com conhecimentos básicos de análise, mas que não conhecem sobre *machine learning*; e (ii) cientistas de dados, ou seja, pessoas que têm o domínio sobre modelos de *machine learning* e podem utilizar a plataforma para acelerar o desenvolvimento de algoritmos e sua implantação em produção. Por ser um projeto voltado ao agronegócio, um dos tipos de perfil poderia vir de *start-ups*, já que esse negócio cresceu consideravelmente nos últimos anos. Esse público, na sua maioria, pertence ao perfil (i) e, em menor parte, ao perfil (ii).

Além disso, a plataforma de IA desenvolvida pode ser aplicada não apenas no agronegócio, mas em qualquer situação. No entanto, ela tenderá a ter mais atratividade para o perfil (i) quanto mais algoritmos estiverem disponíveis e ela terá mais algoritmos disponíveis se tiver adeptos do perfil (ii) que disponibilizem mais algoritmos na plataforma. Isso promove um ciclo virtuoso de adesão, mas que precisa de incentivos exógenos no início do seu ciclo de vida.

A adesão por parte de cientistas de dados pode, em primeiro momento, aumentar a quantidade de algoritmos úteis e reutilizáveis e, em segundo momento, até mesmo aqueles com o perfil (i) começarem a incluir análises estatísticas ou até mesmo automatizar partes do processo das atividades. Mas como incentivar

esse perfil de profissional a se apropriar de tal ferramenta?

A partir de então, deu-se início a criação de um jogo que pudesse instigar os participantes a aprender os primeiros passos dessa plataforma; espera-se que, futuramente, essa experiência possa ser disseminada para outros. O processo de criação do jogo é descrito na próxima seção. Posteriormente, será apresentado o desenvolvimento do conteúdo do jogo, incluindo a base simulada e as etapas de análise e algumas informações da aplicação do jogo. Por fim, algumas considerações são traçadas na seção da conclusão deste artigo.

PROCESSO DE CRIAÇÃO

Uma das grandes dificuldades em entender como montar uma abordagem com potencial de trazer sucesso ao objetivo é justamente captar a essência das lacunas que precisam ser preenchidas para o público-alvo. Por se tratar de uma plataforma que se destina a duas categorias de público-alvo, o conteúdo a ser tratado no jogo deveria ser capaz de atender aos dois perfis.

De maneira geral, qualquer material de apoio ao ensino deve ter conteúdos que puxem as pessoas com menor bagagem de conhecimento no tema para seguirem sem grandes dificuldades com o conteúdo mais avançado. Assim, o processo de criação do jogo Caçada Inteligente contou com o desenvolvimento de vários pilares que são descritos a seguir.

Temática

A plataforma de IA desenvolvida tem funcionalidades básicas que qualquer usuário iniciante deve conhecer para que ele possa aplicar posteriormente em seus próprios casos de análise. Uma vez retirada a barreira inicial sobre como e por onde começar, espera-se que os usuários estejam aptos a darem continuidade no uso da plataforma por si só. As funcionalidades básicas que foram, então, contempladas na criação do jogo são: criação de

⁴ O CPQD é uma instituição independente, focada na inovação com base nas Tecnologias da Informação e Comunicação (TICs).

projeto; entendimento inicial sobre a área de experimentação, incluindo uma visão geral sobre os algoritmos disponíveis; criação de experimentos e treinamento de modelos; criação de novos algoritmos; e implantação dos modelos de *machine learning*.

Para um usuário do público-alvo mais experiente, não seria necessário criar uma base de conhecimento de *machine learning*. No entanto, a plataforma tem sua utilidade para aqueles sem grandes experiências e, portanto, foi incluída na temática a sistematização do processo de modelagem com orientação cadenciada, ou seja, instruções bem específicas do que o usuário deveria realizar para cumprir uma etapa do jogo.

Com isso, cada jogador passaria pelos seguintes temas de *machine learning*: análise descritiva, modelos não supervisionados e modelos supervisionados (regressão e classificação). Além disso, os jogadores seriam direcionados a analisar os resultados dos modelos pelos indicadores e gráficos. Em casos pontuais de dúvidas sobre um modelo ou um resultado em gráfico ou indicadores, os usuários poderiam acessar os tutoriais disponíveis na página do projeto.

Ambientação

Por ser um projeto voltado ao desenvolvimento de uma plataforma de IA para apoio ao agronegócio, seria interessante que o cenário a ser escolhido para ambientar o jogo também estivesse alinhado a esse tema. Com isso, optou-se pelo desenvolvimento de um caso fictício na área da apicultura.

A partir desse núcleo, foi desenvolvido um entorno que pudesse envolver os participantes para que eles tivessem uma motivação a mais para seguir o jogo em suas diversas etapas. Nesse entorno foi adicionado uma porção de mistério, urgência e competição, criando curiosidade nos participantes. O texto abaixo foi apresentado como introdução do jogo.

Livádia, uma comunidade às vésperas de um grande evento!

A grande Festa do Mel é uma feira anual que atrai mercadores de todo o continente em busca de matéria-prima

para a produção de doces e hidromel. Este grande evento é o responsável pela prosperidade da comunidade, que depende desta semana para prover o sustento do ano inteiro a toda a sua população.

A comunidade Livádia estava se preparando para o grande evento. Os fazendeiros produziram toneladas de mel para suprir todas as negociações com os mercadores durante a Feira.

Mangueirinha, Jardim d'Oeste, Morro da Colmeia e Vereda Tropical são as fazendas de produção de mel de Livádia. Elas produzem mel de eucalipto, flor de laranjeira, floresta, urze, flores silvestres e girassol.

Após a colheita do mel, ele segue para beneficiamento e envase por quatro armazéns: Avohai, Só Mel, Mel e Cia e Benemel. O mel é envasado e enviado às tendas para comércio.

A comunidade está em polvorosa!

Apenas cinco dias antes do grande evento, um membro da comunidade, que tinha adquirido um pote de mel, sentiu um sabor pungente no produto, o que poderia ser um desastre para toda a comunidade, caso os mercadores se recusassem a negociar a produção do ano.

Apenas algumas horas depois, outros dois casos idênticos ocorreram com membros da comunidade sem qualquer grau de parentesco ou proximidade. Não demorou muito para outros cinco casos aparecessem, ao cair da noite.

A comunidade, que atribuiu o ocorrido à feitiçaria, se uniu para pedir ajuda à Melissa, Deusa da campina e do mel, para descobrir o que houve e salvar o grande evento.

A Deusa enviou caçadores com habilidades sobrenaturais capazes de analisar, entender e salvar a Festa anual do mel. Aqueles que conseguissem resolver o enigma do mel que fora enfeitado para evitar sua comercialização e para destruir a reputação da comunidade ganharia a recompensa.

Os caçadores devem ficar atentos às pistas que aparecem ao longo da jornada!

Esse texto introduziu os participantes ao tema juntamente com uma figura interativa que mostrava um mapa com acesso ao tutorial da

plataforma, ao vídeo de treinamento da plataforma, ao site do jogo e a essa introdução. Além disso, o mapa escondia pistas que poderiam ajudar ou serem ineficazes na busca pela resposta. A Figura 1 mostra o mapa desenvolvido para o jogo.

Figura 1 – Mapa da comunidade Livídia



* Ao passar o mouse pelas peças, os participantes encontram pistas e informações.
Fonte: acervo dos autores.

Foram fornecidas sete pistas referentes à ambientação e uma pista referente aos desafios do jogo. Seis, das sete pistas, foram lançadas no mapa logo na sua introdução aos participantes. Uma delas foi introduzida no último dia do jogo junto com a pista referente aos desafios. A motivação para essa decisão é descrita na subseção Cadência de ensino. As pistas que foram colocadas no mapa são:

- erros humanos podem ocorrer;
- esporos de fungos e bactérias podem permanecer no ar por longos períodos de tempo e serem deslocados por grandes distâncias, além de serem resistentes ao calor, agentes químicos e físicos;
- anos antes, a Feira anual de trigo de uma comunidade sofreu sabotagem fazendo com que toda a produção fosse perdida;
- meses antes, membros da comunidade relataram que viram uma carruagem que levava o mel de Mangueirinha para Avohai parada e sem a presença do cocheiro;

- fungos podem ser encontrados no intestino de abelhas operárias e, ocasionalmente, levam ao adoecimento da colmeia;
- a quantidade de alcaloides abaixo de 1.000 (mil) não afeta, necessariamente, as características nutricionais e de sabor;
- do desafio 6 em diante, se errar, é como pular uma casa do botão.

Com o tema introduzido, segue-se, então para as regras do jogo. Nessa versão, o jogo se dá em um evento de 4 a 5 dias, sendo os quatro primeiros voltados ao cumprimento dos desafios com apoio de equipe especialista e o último voltado às últimas revisões e consolidação de pontos por parte da organização. Essa definição foi motivada pelo fato de que os participantes não poderiam dispor de períodos completos para a realização das atividades. Assim, foi proposto que o jogo levasse mais dias e menos tempo por dia. No entanto, em situações diferentes, o jogo poderia levar apenas um dia ou algumas horas.

Ao iniciar uma sessão de uso da plataforma, os jogadores receberiam o primeiro de 10 desafios. Após completar e enviar resposta por meio de um formulário eletrônico, os jogadores receberiam sua pontuação e um novo desafio. A pontuação poderia ser:

- desafio concluído integralmente [Pontuação total];
- desafio com resultados parciais [Pontuação parcial];
- desafio não concluído [Não pontua].

Se a equipe obtiver resultado A, ela recebe o próximo desafio (um novo formulário e, em alguns casos, um arquivo de dados pelo *e-mail*). Se a equipe obtiver o resultado B ou C, ela tem a opção de retornar para completar o desafio e receber a respectiva pontuação ou de seguir em frente com a pontuação obtida. Nesse último caso, ela recebe o próximo desafio (um novo formulário e, em alguns casos, um arquivo de dados pelo *e-mail*). A chance de a equipe voltar no mesmo desafio e tentar completá-lo corretamente tem o objetivo de incentivar o aprendizado sem ônus de pontuação, mantendo-a no páreo da competição. Além disso, para cada desafio, as equipes poderiam:

- pedir uma dica sem custo na pontuação;
- pedir uma dica adicional com um pequeno custo na pontuação;
- pedir 15 minutos de consultoria com especialistas a um custo maior na pontuação.

Como um dos públicos-alvo da plataforma são pessoas sem conhecimento na área de *machine learning*, uma forma de manter os participantes interessados no jogo foi criar desafios bônus que tivessem um nível de dificuldade baixo e garantissem uma reposição de pontos que pudessem ter sido descontados por não terem sido completados os desafios principais corretamente.

Critérios de avaliação

Cada desafio poderia ter de duas a cinco atividades. Foi desenvolvido um quadro de pontuação que tinha alguma relação com o nível de dificuldade e de importância para a resposta final do jogo. A pontuação detalhada por desafio e por atividade está apresentada na Tabela 1.

Tabela 1 – Pontuação dos desafios por atividade

Desafios	Atividades					
	1	2	3	4	5	6
D1	100	100	100	100	100	500
D2	100	400	-	-	-	500
D3	500	500	800	-	-	1800
D4	100	500	800	-	-	1400
D5	300	300	-	-	-	600
D6	500	800	800	-	-	2100
D7	100	300	300	800	-	1500
D8	300	300	-	-	-	600
D9	500	500	-	-	-	1000
D10	2000	1500	1000	500	-	5000

Fonte: elaborada pelos autores.

Os pontos por atividade poderiam ser divididos em quatro faixas: 0%, 30%, 70% ou 100%. Além disso, os desafios bônus eram pontuados da seguinte forma: até 500 pontos para o desafio A, até 700 pontos para o desafio B e até 1200 pontos para o desafio C. Apenas as equipes que estivessem com suas pontuações incompletas poderiam pontuar no desafio bônus. A pontuação do bônus nunca

ultrapassaria o limite de pontuação total dos desafios (15.000).

DESENVOLVIMENTO DO CONTEÚDO

Todo o desenvolvimento do conteúdo do jogo foi inspirado nas relações reais sobre o tema da apicultura. No entanto, evitou-se

manter uma correlação perfeita para que o jogo pudesse incentivar a análise crítica e não uma busca por informações na internet. As subseções a seguir descrevem a abordagem do jogo e a criação da base de dados simulada.

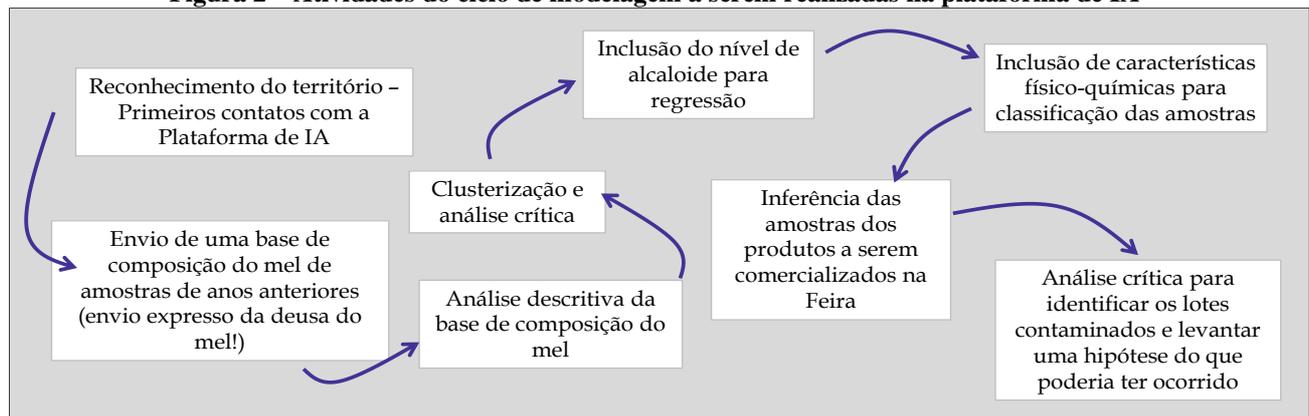
Cadência de ensino

O primeiro passo para o jogo foi o de criar um tutorial em formato de vídeo de aproximadamente 20 minutos, sem narração, como um cinema mudo. Esse vídeo foi postado no *YouTube* e no site do evento (o vídeo pode ser acessado pelo link: https://www.youtube.com/watch?v=PL__RaojxAg). O vídeo é iniciado com um resumo simplificado sobre o ciclo de desenvolvimento de um modelo de *machine learning* por meio de

figuras e poucos textos. O objetivo do vídeo era comunicar rapidamente que tipo de ações a plataforma se propõe a facilitar, tendo em vista que o público-alvo menos experiente não conhece sobre o tema. Após essas primeiras cenas referentes ao ciclo de modelagem, o vídeo apresenta a plataforma de IA desenvolvida ao longo do projeto, passeando pelas suas funcionalidades elementares. O vídeo foi disponibilizado quatro dias antes do início do evento Caçada Inteligente.

Os desafios da Caçada Inteligente estavam diretamente associados às funcionalidades elementares da Plataforma e aos pontos do ciclo de modelagem apresentados no treinamento. A Figura 2 mostra as etapas que foram tratadas em forma de 10 desafios.

Figura 2 – Atividades do ciclo de modelagem a serem realizadas na plataforma de IA



Fonte: elaborada pelos autores.

O primeiro desafio estava associado apenas a direcionar o participante nos primeiros passos da plataforma, tirando *prints* de partes específicas de tutoriais e de pequenas ações a serem realizadas, como criar um novo projeto na plataforma e renomear um experimento. No segundo desafio, as equipes recebiam, além do formulário de resposta com o respectivo desafio, um arquivo no formato csv que continha uma massa de dados fictícia sobre a composição do mel. As tarefas que compunham esse desafio são: acessar o tutorial sobre o modelo não supervisionado *k-means* e realizar uma análise descritiva dos dados por meio de um algoritmo da plataforma.

O terceiro desafio contempla o treinamento de dois modelos a partir do algoritmo *k-means*, sendo um deles com a padronização dos dados

e o outro com os dados originais. Em seguida, pedia-se que fosse realizada uma análise descritiva agrupando por cluster. O objetivo era direcionar o participante a realizar uma análise crítica dos resultados como uma forma de treiná-lo a fazer o mesmo para, ao longo dos desafios, ele ter o discernimento de dar continuidade nesse método para identificar o problema ocorrido no mel.

O quarto, quinto e sexto desafios foram relacionados ao treinamento de modelos de regressão; porém, no meio do caminho, os participantes teriam que identificar uma inconsistência nos dados enviados em um novo arquivo. Esse novo arquivo trazia não apenas os dados já enviados, mas também os rótulos para treinamento da regressão. Entre os rótulos era possível observar valores negativos em uma

variável que só poderia conter valores maiores ou iguais a zero.

Uma das atividades do desafio 5 era criar um novo algoritmo que fizesse a exclusão de linhas de uma base a partir de uma condição. Como os participantes não tinham conhecimento de programação, todos os passos para a criação desse algoritmo para incluí-lo na plataforma foram descritos detalhadamente, inclusive as linhas de código que deveriam ser copiadas no editor de códigos da plataforma (JupyterLab). Após a criação desse algoritmo, ele deveria ser encadeado com a leitura de dados e a nova base poderia ser baixada para, então, realizar o treinamento de regressão sem os dados problemáticos.

Assim, no sexto desafio, os participantes testaram vários tipos de modelos de regressão, cujos algoritmos estão disponíveis na plataforma para, em seguida, comparar os resultados em tela única e eleger o melhor modelo. Para eleger o melhor modelo, os participantes foram direcionados a um gráfico específico que mostrava a distribuição do erro e fornecia indicadores em percentual (medida que é mais conhecida pelos não especialistas).

O sétimo e oitavo desafios foram relacionados ao treinamento de modelos de classificação. Com isso, todos receberam uma nova base que continha todos os dados dos desafios anteriores, inclusive o rótulo da regressão, mais os rótulos para treinamento de modelos de classificação. Os rótulos de regressão, nessa fase, passaram a fazer parte das variáveis explicativas do modelo de classificação. Como resultado, os participantes eram direcionados a avaliar a matriz de confusão. Assim como no caso do gráfico de distribuição de erro na regressão, essa é uma forma bastante direta de avaliar resultados dos modelos treinados. Ainda assim, tanto no caso da regressão como no caso da classificação, a primeira atividade dos desafios foi de acessar o tutorial e tirar um *print* desses artefatos. Com isso, o próprio participante poderia ler sobre como interpretar esses indicadores sem a ajuda de dicas ou de consultoria, que custariam alguns pontos na competição.

Para o desafio 9 os participantes recebiam uma nova base. Suas atividades eram utilizar a funcionalidade de pré-implantação de modelos

e testar os resultados com essa base exclusiva de inferência. Nela, apenas dados das variáveis iniciais eram fornecidas. Com isso, era esperado que os participantes utilizassem o modelo implantado de regressão para estimarem o valor do respectivo rótulo, baixassem esse resultado e utilizassem essa nova base com as estimativas de regressão para realizar a inferência do modelo de classificação. A base resultante é o material a ser utilizado para identificar os lotes impróprios de mel, no desafio 10. Esse último desafio requeria que os participantes não só apontassem os lotes impróprios, mas também descrevessem como chegaram à conclusão e lançassem uma hipótese sobre o motivo do problema. Por fim, era solicitado que os participantes implantassem os modelos para futuras análises fictícias.

Tema de análise

O tema escolhido para a criação do jogo foi o de produção de mel. A partir desse momento, iniciou-se uma etapa de busca por informações sobre essa área que pudessem servir de inspiração para a criação de uma base de dados fictícia, porém baseada em inspirações do comportamento real desse negócio.

O primeiro passo foi pesquisar sobre a composição básica do mel e sobre indicadores comuns de avaliação de qualidade. Não se esperava, contudo, uma busca exaustiva por uma massa de dados reais por dois motivos principais. O primeiro deles é que uma base real poderia fazer com que os participantes encontrassem sua análise em buscas na internet, distorcendo o foco do jogo. O segundo e mais importante motivo é que, com uma base simulada, é possível criar relações propositalmente para que o jogo siga o rumo esperado, passando por todas as etapas para o aprendizado dos usuários potenciais.

Em Anacleto e Marchini (2004) e em Suarez et al. (2010) foram obtidas informações gerais, como percentual de glicose, sacarose, proteínas, acidez, pH etc. Considerando essas informações básicas como princípio de uma base simulada, partiu-se para uma busca do comportamento que as colmeias poderiam ter em relação às doenças prováveis que poderiam

diminuir ou comprometer a qualidade do mel. Em Medeiros e Souza (2015), por exemplo, os autores discutem sobre aspectos da contaminação do mel, inclusive mencionando contaminação por metais pesados e antibióticos. A contaminação microbiana também foi mencionada, sendo esta a mais discutida. Nesse assunto, os textos visitados levavam a um ponto interessante: a disponibilidade de água, que poderia aumentar as chances de contaminação e o pH (FLORES et al., 2000). Outro ponto levantado foi o da produção de toxinas a partir de contaminação por fungos ou bactérias e seus esporos. Em Lanza (2016) é mencionado um problema da

toxina alcaloides pirrolizidínicos e seus níveis de tolerância no mel.

Após alguns estudos para poder criar uma base de dados que tivesse elementos reais da natureza, mas que incluíssem a dose necessária de ficção, deu-se origem às etapas para a montagem dessa base fictícia. Essa montagem não levou em conta as possibilidades reais de doenças, mas manipulou as relações para criar um novo contexto natural que pudesse fazer sentido ao analista. A Tabela 2 mostra os fatores Fazenda, Cultura e Armazéns de envase e suas combinações existentes na base.

Tabela 2 – Combinações de Fazenda, Cultura e Armazém de envase encontradas na base

Fazenda	Cultura	Armazéns de Envase			
Mangueirinha	Eucalipto	Avohai	Benemel	Mel e Cia	Só Mel
Mangueirinha	Floresta				
Jardim D'Oeste	Girassol				
Morro da Colmeia	Urze				
Vereda Tropical	Flor de Laranjeira				
Vereda Tropical	Flores Silvestres				

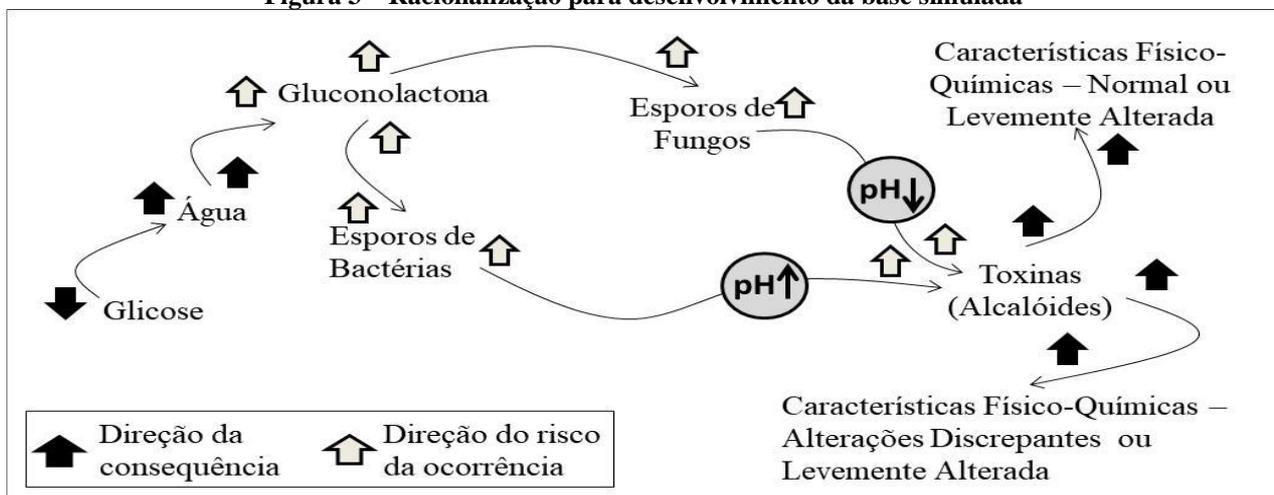
Fonte: elaborada pelos autores.

Na Tabela 2 é possível observar que duas das fazendas produzem mel a partir de duas culturas distintas cada e duas delas produzem mel a partir de uma cultura única. Todas as fazendas e culturas são enviadas para envase em qualquer um dos quatro armazéns.

Essas combinações ganharam características relativamente distintas uma da

outra, mas sem que fossem diretamente relacionadas às duas principais variáveis do jogo: o nível de alcaloide (toxina) para o modelo de regressão e a característica físico-química para o modelo de classificação. A Figura 3 mostra a racionalização para a construção da base simulada.

Figura 3 – Racionalização para desenvolvimento da base simulada



Fonte: elaborada pelos autores.

Com o levantamento das variáveis iniciais, foram gerados valores aleatórios com a distribuição normal, cuja média e desvio padrão foram obtidos de Anacleto e Marchini (2004) na maioria dos casos. Seguiu-se, então, com distorções ou deslocamentos dos valores por grupos, por exemplo: Fazenda, Fazenda com Cultura etc. Esses valores das variáveis iniciais, como glicose, sacarose, maltose, minerais, índice de formol etc. eram tomados como referência para estimar as variáveis que poderiam carregar algum grau de dependência (valor absoluto da correlação abaixo de 0,7). Assim, como mostra a Figura 3, um nível de glicose mais baixo implica que o nível de água estaria mais alto. Com mais água, há maior probabilidade de aumento do nível de gluconolactona que, por sua vez, aumenta o risco de contaminação por bactérias e fungos. No caso de haver uma contaminação por esporos de fungos, se o pH do meio estiver baixo, aumenta a probabilidade de formação de toxinas (no jogo, chamado de alcaloide). Dependendo do nível dessas toxinas produzidas a partir de fungos, as características físico-químicas continuam normais ou levemente alteradas. No entanto, no caso de haver uma contaminação por bactérias ou seus esporos, se o pH estiver alto, maior é o risco de formação de toxinas e, nesse caso, maior a chance de alterações discrepantes ou leves nas características físico-químicas.

Foram introduzidas, também, algumas relações do tipo: cinzas e minerais, acidez e ácido glucônico e agrupamentos de valores por cultura. Também foi introduzido um nível de metais pesados, nos Armazéns de envase Só mel e Avohai, e de antibióticos, apenas na Fazenda Mangueirinha. No entanto, essas relações não estavam correlacionadas com o rumo principal da análise, mas sim com uma forma de adicionar nuances à base.

Parte dessa base foi separada para ser utilizada como base para inferência. Ou seja, a base de treinamento de modelos foi classificada como uma base de análises de anos anteriores e que só serviriam para dar conhecimento aos participantes do jogo sobre os padrões. No entanto, a resposta para que os lotes de mel pudessem ser apontados seria proveniente da base de inferência, que foi classificada como

amostras dos lotes de mel atuais em que estavam contidos os produtos estragados. Essa base de inferência foi filtrada e teve os lotes selecionados de modo que apenas alguns casos de alterações discrepantes e levemente alterados estivessem presentes. Além disso, foi selecionado apenas um caminho que levava à resposta correta: alto pH, alto nível de alcaloide e alterações discrepantes.

Realizando todos os desafios e fazendo as análises críticas dos resultados era possível apontar que os lotes que deveriam ser eliminados da festa pertenceriam à Fazenda Vereda Tropical, da cultura Flor de laranjeira e envasado pelo armazém Mel e Cia que tinham o pH alto.

APLICAÇÃO DO JOGO

O jogo, chamado Caçada Inteligente, é voltado para interessados em conhecer e aprender a utilizar as funcionalidades da plataforma de IA de maneira rápida e atrativa. Uma primeira aplicação do jogo foi realizada com colaboradores da empresa desenvolvedora e, posteriormente, com pessoas voltadas a análises do agronegócio. Inicialmente, a expectativa era de que aproximadamente 20 duplas se inscrevessem; no entanto, em três dias de inscrições, a organização do evento recebeu 50 duplas inscritas. O evento teve início no dia 12 de julho de 2021 e finalizou em 16 de julho de 2021. Das 50 duplas, 35 de fato iniciaram o jogo respondendo ao primeiro desafio, 29 no segundo desafio, seguindo sequencialmente por 28, 23, 21, 19, 19, 15 e, no décimo e último desafio, 13 duplas enviaram suas respostas. Poucas equipes utilizaram a facilidade da consultoria e das dicas adicionais (que tinham custo na pontuação). Por sua vez, a dica gratuita teve muitos adeptos.

De maneira geral, os participantes não tiveram dificuldades no uso da tecnologia; no entanto, apesar das instruções, tutoriais e do vídeo de treinamento, houve uma dificuldade relatada por vários quanto à análise crítica dos dados. Em uma avaliação realizada após o evento, por meio de um questionário no *Google Forms*, foram obtidas 23 respostas de participantes. Destes, a maioria avançou nos desafios. No entanto, entre os que abandonaram

o jogo, 90% dos respondentes disseram que foi por falta de tempo disponível por conta das atividades do trabalho. De qualquer forma, a principal motivação para a participação no evento foi aprender algo sobre IA, em primeiro lugar, e, em segundo lugar, conhecer a plataforma. Ao serem perguntados sobre o vídeo de treinamento da plataforma, 52% mencionaram que deveria ter explicação falada e legendada, 40% acharam o vídeo adequado e os outros 8% acreditam que seria necessário ser mais detalhado. Quanto ao grau de dificuldade dos desafios, quase 80% dos participantes consideraram que eram difíceis ou muito difíceis. Essa era uma resposta esperada, visto que não era um público com conhecimento prévio em *machine learning*. Sobre o jogo Caçada Inteligente, 35% dos participantes que responderam ao formulário deram nota 10 para o evento, 70% deram nota acima de 8, 22% notas 6 e 7 e apenas dois casos deram nota 3. Por fim, ao serem perguntados sobre a plataforma desenvolvida, 95% deram nota acima de 8 e apenas um caso (5%) deu nota 7 para a plataforma, mostrando uma boa aceitação da ferramenta. Apesar das dificuldades em geral quanto à interpretação dos resultados dos desafios, duas equipes conseguiram chegar à resposta correta, apontando os lotes que deveriam ser excluídos da festa e salvando a comunidade.

CONCLUSÕES

O formato do jogo foi estimulante para os participantes e pode ser aplicado para concentrar o aprendizado no mesmo tempo para muitas pessoas. No entanto, para um formato de aprendizado sob demanda, seria necessário realizar alguns ajustes. Assim, para que essa ideia seja adequada para um formato do tipo *wizard*, as respostas aos desafios devem ser automatizadas e as dicas poderiam ser num formato de assistente virtual simplificado.

De forma geral, pôde-se observar um engajamento forte daqueles que passaram da metade dos desafios, inclusive demandando informações e acesso à plataforma para aprofundamento futuro.

Agradecimentos

Este projeto contou com o apoio do Ministério da Ciência, Tecnologia e Inovação e da Rede Nacional de Ensino e Pesquisa).

REFERÊNCIAS

- ANACLETO, D.; MARCHINI, L. Composição físico-química de amostras de méis de *Apis mellifera* L. provenientes do cessado paulista. **B. Indústr. anim.**, Nova Odessa, v. 61, n. 2, p.161-172, 2004.
- FLORES, J. et al. Ascosferiose (*Ascosphaera apis*): Causas predisponentes, medidas de controle e prevenção. **B. Indústr. anim.**, Nova Odessa, v. 57, n. 2, p.201-209, 2000.
- LANZA, J. Mel e o risco de toxinas originadas das plantas de origem. **FoodSafetyBrazil.org**. 2 out. 2016. Disponível em: <https://foodsafetybrazil.org/mel-e-o-risco-de-toxinas-originadas-das-plantas-de-origem>. Acesso em: 09 jul. 2021.
- MEDEIROS, F.; SOUZA, M. Contaminação do mel: a importância do controle de qualidade e de boas práticas apícolas. **Atas de Ciências da Saúde**, v. 3, n. 4, 2015.
- SUAREZ, J. M. et al. Contribution of honey in nutrition and human health: a review. **Mediterranean Journal of Nutrition and Metabolism**, v. 3, n. 1, p. 15-23, 2010.

DADOS BIOGRÁFICOS DOS AUTORES



Graziella Cardoso Bonadia – Possui graduação em Estatística pelo Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas (2000), mestrado em Engenharia Elétrica pela Universidade Estadual de Campinas (2019). Atualmente trabalha no CPQD e é doutoranda em Engenharia Elétrica pela Universidade Estadual de Campinas. Tem experiência na área de Inteligência Artificial, atuando em projetos de pesquisa e desenvolvimento com ênfase em Processamento de Linguagem Natural. Foi líder técnica do projeto de desenvolvimento da plataforma de IA.



Fabiani de Souza – Possui graduação em Ciência da Computação pelo Instituto de Computação da Universidade Estadual de Campinas (2014). Atualmente trabalha no CPQD e cursa disciplinas como aluna especial na pós-graduação da Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas. Tem experiência na área de Ciência da Computação, com ênfase em Interação Humano-Computador, atuando principalmente nos seguintes temas: experiência do usuário, usabilidade, acessibilidade e inclusão digital.



Gilmar Barreto – Possui graduação em Engenharia Química pela Universidade Estadual de Campinas (1982), mestrado em Engenharia Elétrica pela Universidade Estadual de Campinas (1986) e doutorado em Engenharia Elétrica pela Universidade Estadual de Campinas (2002). Atualmente é professor da Universidade Estadual de Campinas. Tem experiência na área de Engenharia Elétrica, com ênfase em Eletrônica Industrial, Sistemas e Controles Eletrônicos, atuando principalmente nos seguintes temas: sistemas nebulosos, sistemas multivariáveis, controle, otimização multiobjetiva e eletroquímica tem atuado intensamente na área de Ensino de Engenharia. É autor do livro “Veículos Elétricos” em coautoria com o Prof. Dr. Celso Pascoli Bottura e do livro “Circuitos de Corrente Alternada” em coautoria com os professores Carlos Alberto de Castro Junior, Carlos Alberto Favarin Murari e Fujio Sato.