

## Informativo 02

São Paulo, 22 de setembro de 2023.

### **29ª Competição Nacional Baja SAE BRASIL** **Relatório de Projeto - Desafio Machine Learning – REVISÃO 2**

**Contexto:** A inteligência artificial é um dos catalisadores do desenvolvimento tecnológico em um mundo cada vez mais automatizado e digital. O seu verdadeiro potencial está em sua aplicação à tomada de decisão durante o design, construção e operação de ativos. A tomada de decisão baseada em dados pode ajudar a prever resultados e evitar erros durante o projeto, além de garantir performances otimizadas visando o melhor desempenho dos produtos.

**Demanda/Entrega:** O objetivo deste desafio é fazer com que a equipe desenvolva modelos de *Machine Learning* (ML) da performance dos carros nos eventos dinâmicos da competição e, a partir desses, prever a pontuação que sua equipe irá alcançar nas avaliações dinâmicas da Competição Baja SAE BRASIL 2024 – Etapa Nacional.

A fim de auxiliar na construção de modelos que retratem a realidade da competição, as equipes terão acesso a um banco de dados com a informação histórica de veículos e suas respectivas pontuações em provas de competições passadas através de um *software* de inteligência artificial, a ser descrito mais à frente. **Como bem sabemos, tais informações são confidenciais, portanto o banco de dados não revelará de quais equipes/veículos são os resultados, nem mesmo o ano e evento de participação.**

Após o desenvolvimento dos modelos ML, as equipes utilizarão os dados do próprio veículo, preenchidos na ficha técnica apresentada para competir na 29ª Competição BAJA SAE BRASIL 2024 – Etapa Nacional, para prever quantos pontos irão obter nos eventos dinâmicos. As equipes deverão apresentar as respectivas informações conforme o exemplo da Tabela 1. **É de fundamental importância apresentar os resultados neste formato, de forma a padronizar a correção dos desafios. A equipe que modificar este formato será penalizada.**

	Aceleração	Velocidade	Frenagem	Tração	Manobrabilidade	Suspensão
Desempenho do Modelo ML	82%	89%	88%	97%	64%	67%
Grupo Previsto	4	3	5	4	2	3

**Tabela 1:** Exemplo de tabela que deverá ser apresentada neste desafio.

No relatório do desafio, a equipe deverá descrever todas as etapas realizadas no desenvolvimento dos modelos ML. Os tópicos básicos e **OBRIGATÓRIOS** a serem apresentados são:

- Limpeza e Pré-Processamento dos dados: nas aplicações de ML, a limpeza dos dados é pré-requisito para a construção de modelos fidedignos, pois os dados brutos podem conter erros (dados incorretos, incompletos, duplicados, fora de formatação etc.) que afetem a precisão da modelagem, provocando previsões errôneas que afetem negativamente o resultado. Além disso, a análise crítica dos dados faz com que a equipe desenvolva capacidade interpretativa dos resultados e compreenda a correlação de informações, competências necessárias para resolução do desafio, que fornecerá um banco de dados contendo tanto dados reais quanto dados “sujos”. Por fim, como existem diversos métodos para análise e limpeza de dados, a equipe deverá descrever a metodologia escolhida, bem como apresentar as referências bibliográficas utilizadas para tal no modo **Estilo Numerado (Estilo 1)**;
- Escolha do modelo: um modelo de *Machine Learning* é um programa que descobre os padrões e extrai decisões com base na base de dados escolhida e tratada. Eles podem ser treinados para executar diversos tipos de funções. Por exemplo, temos modelos de regressão linear, árvores de decisão, *clustering* etc. Para desenvolvimento deste tópico, a equipe deverá entender os tipos de algoritmos envolvidos em ML, de modo a escolher o modelo que mais se aproxima dos objetivos do Desafio Técnico. Dentro de algumas ferramentas comerciais, é possível criar modelos ML de forma automática (*Auto Models*), o que facilita a escolha do tipo de modelo apropriado para cada *dataset*. A equipe deverá descrever a escolha do modelo, bem como apresentar as referências bibliográficas utilizadas para tal **Estilo Numerado (Estilo 1)**;
- Treinamento do modelo: o treinamento do modelo consiste em executar o algoritmo desenvolvido em um conjunto de dados de forma a otimizá-lo e entender padrões e resultados oriundos da modelagem. A equipe deverá descrever o processo de treinamento do modelo, descrevendo a base de treino utilizada, bem como as justificativas para a escolha;
- Avaliação do modelo: para avaliação do modelo, se ele de fato aprendeu o algoritmo, a equipe deverá utilizar métricas de forma a avaliar criticamente e quantitativamente a qualidade do modelo escolhido. Essas métricas deverão ser descritas e justificadas;
- Aprimoramento dos parâmetros: nesta etapa, a equipe deverá ser capaz de identificar valores que afetam diretamente a precisão do modelo. A equipe deverá justificar a metodologia utilizada para aprimoramento;
- Predição: Por fim, com base no modelo detalhado e aprimorado nas etapas anteriores, apresente a predição das pontuações a partir dos dados técnicos do protótipo da sua equipe.

## Observações:

1. A previsão de pontuação poderá exigir um modelo e sistema de aprimoramento específico para cada uma das provas dinâmicas. No relatório será necessário descrever todas as etapas básicas de **apenas um** dos modelos e, para os demais, basta explicitar e justificar seus diferenciais em relação ao modelo base;
2. Não esquecer de apresentar as referências tais como: dados oficiais, normas, artigos científicos, publicações em revistas e livros ou qualquer outra fonte desde que seja reconhecida no meio acadêmico;
3. A pontuação do Desafio Técnico será dividida em duas partes:
  - **70%** - Avaliação técnica do conteúdo apresentado no relatório pelo corpo de Juízes.
  - **30%** - Análise comparativa da Pontuação Prevista vs. Pontuação Real, obtida durante a 29ª Competição BAJA SAE BRASIL 2024 – Etapa Nacional.
4. Referente a Análise Comparativa, a pontuação será calculada conforme procedimento descrito a seguir.
  - Tanto as pontuações previstas, quanto as reais (alcançadas na competição), deverão ser classificadas em grupos de acordo com as faixas pré-definidas na Tabela 2.

	Pontuação máxima por prova	Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
Aceleração	30	0-6	7-12	13-18	19-24	25-30
Velocidade	30	0-6	7-12	13-18	19-24	25-30
Frenagem	20	0-4	5-8	9-12	13-16	17-20
Tração	50	0-10	11-20	21-30	31-40	41-50
Manobrabilidade	50	0-10	11-20	21-30	31-40	41-50
Suspensão	70	0-14	15-28	29-42	43-56	57-70

**Tabela 2:** Classificação dos grupos para as pontuações previstas.

- Tomando como exemplo a “Equipe X”, para cada prova dinâmica uma pontuação foi prevista (por meio dos trabalhos realizados e apresentados no relatório do Desafio Técnico), e uma pontuação foi alcançada (por meio do desempenho na competição). Na Tabela 3 é possível observar as pontuações hipotéticas e os respectivos grupos de classificação, conforme faixas da Tabela 2.

Exemplo Equipe X	Pontuação Prevista	Grupo Previsto	Pontuação Real	Grupo Real
Aceleração	24	4	25	5
Velocidade	18	3	15	3
Frenagem	17	5	5	2
Tração	39	4	43	5
Manobrabilidade	12	2	27	3
Suspensão	36	3	57	5

**Tabela 3:** Exemplo da pontuação vs. Grupo, conforme Tabela 2.

- A correlação entre o Grupo Previsto e o Grupo Real é apresentada na grade da Tabela 4, onde para cada combinação um nível de predição é definido – sendo eles: Acerto, Erro Médio e Erro Alto.

		Grupo Previsto				
		Grupo 1	Grupo 2	Grupo 3	Grupo 4	Grupo 5
Grupo Real	Grupo 1	Acerto	Erro Médio	Erro Alto	Erro Alto	Erro Alto
	Grupo 2	Erro Médio	Acerto	Erro Médio	Erro Alto	Erro Alto
	Grupo 3	Erro Alto	Erro Médio	Acerto	Erro Médio	Erro Alto
	Grupo 4	Erro Alto	Erro Alto	Erro Médio	Acerto	Erro Médio
	Grupo 5	Erro Alto	Erro Alto	Erro Alto	Erro Médio	Acerto

**Tabela 4:** Grade de classificação da comparação Previsto vs. Real

- De forma complementar a Tabela 4, para cada nível de predição uma pontuação é atribuída (ver Tabela 5). Sendo a pontuação máxima de 30% alcançada com o “Acerto” em todas as 6 provas dinâmicas.

	Erro Alto	Erro Médio	Acerto
Pontuação por cenário	0.0%	2.5%	5.0%

**Tabela 5:** Critério para as notas proporcionais.

- Dando continuidade ao exemplo da “Equipe X”, para cada correlação do Grupo Previsto vs. Grupo Real, um nível de predição foi definido com base na Tabela 4. Por consequência, uma pontuação foi calculada para cada prova com base na Tabela 5. Ver Tabela 6.

Exemplo Equipe X	Grupo Previsto	Grupo Real	Nível de previsão	Pontuação por prova
Aceleração	4	5	Erro Médio	2.5%
Velocidade	3	3	Acerto	5.0%
Frenagem	5	2	Erro Alto	0.0%
Tração	4	5	Erro Médio	2.5%
Manobrabilidade	2	3	Erro Médio	2.5%
Suspensão	3	5	Erro Alto	0.0%
<b>Soma Total:</b>				<b>12.5%</b>

**Tabela 6:** Continuação do exemplo com cálculo da pontuação

- Por fim, no exemplo apresentado na Tabela 6, a “Equipe X” obteve a soma total de 12.5% do total de 30.0%, referente a segunda parte da nota do Desafio Técnico.
5. A publicação das notas será em dois momentos: Desafio Técnico – Relatório, acompanhando as notas do relatório de projeto; e Desafio Técnico – Dinâmicas, no sábado de competição após os eventos dinâmicos.
  6. No caso de a equipe não entregar ou ter o relatório do Desafio Técnico desclassificado, será considerado “Erro Alto” para todas as provas, ou seja, 0.0% dos 30.0%. **Visto isso, é de suma importância que as equipes apresentem o Desafio Técnico.**

#### Orientações gerais (#ficammaisdicas):

1. Lembre-se que os juízes não acompanharam o desenvolvimento deste trabalho. Portanto seja claro e objetivo com as informações;
2. As informações contidas neste trabalho deverão ser coerentes com os dados apresentados no relatório de projeto;
3. Todas as previsões apresentadas serão confidenciais e não serão expostas para outras Equipes até os eventos dinâmicos da Competição Baja SAE BRASIL 2024– Etapa Nacional;
4. Caso a Equipe ainda não possua nenhuma ferramenta de ML liberada para uso dentro da instituição de ensino, uma dica é solicitar uma licença estudantil gratuita do programa Altair RapidMiner.
5. Dentre todos os eventos dinâmicos, **APENAS** a prova **SUPER PRIME** foi desconsiderada da avaliação do Desafio Machine Learning.
6. **APENAS** os eventos dinâmicos citados neste informativo (**Aceleração, Velocidade, Frenagem, Tração, Manobrabilidade e Suspensão**) serão considerados na avaliação do Desafio Machine Learning.

7. Para citar de forma correta as referências, siga os procedimentos no seguinte endereço:

<https://booksite.elsevier.com/9780081019375/content/Elsevier%20Standard%20Reference%20Styles.pdf>

#### Patrocinador Oficial do Desafio:

A Altair Engineering do Brasil é a patrocinadora oficial do Desafio Técnico da Competição Nacional Baja SAE BRASIL 2024 – Etapa Nacional.

A Altair é uma empresa de tecnologia que desenvolve soluções nas áreas de Simulação, Inteligência Artificial (IA), *Data Analytics* e Computação de Alto Desempenho.

“A convergência da ciência de dados e da simulação baseada em IA se resume a uma proposta de valor simples, mas poderosa: quanto mais rápido você demonstrar a verdade e/ou o valor de uma ideia ousada, mais rápido você poderá concretizar essa ideia – e as vantagens que a acompanham. Essa mentalidade impulsiona tudo o que fazemos. Nossa visão é transformar a tomada de decisões empresariais.”

Para as Equipes que escolherem a ferramenta de ML da ALTAIR:

1. Enviar um e-mail para [brasil@altair.com](mailto:brasil@altair.com);
2. No e-mail, descrever que irão participar da Competição Baja SAE BRASIL 2024 – Etapa Nacional, e irão realizar o Desafio *Machine Learning*.
3. Todas as Equipes receberão gratuitamente um material de treinamento introdutório de *Data Analytics* e um guia do software RapidMiner. Além disso, receberão um treinamento online (ao vivo) com cientistas de dados especialistas na ferramenta.
4. No dia **29/09/2023 às 20hs** haverá uma transmissão ao vivo com a participação do presidente dos eventos estudantis do BAJA SAE BRASIL, o Comitê Técnico da Competição Baja SAE BRASIL 2024 e os especialistas de dados da ALTAIR. O objetivo deste evento é realizar um alinhamento técnico entre os competidores e o Comitê Técnico do DESAFIO TÉCNICO 2024. Dada a importância e a profundidade do assunto, a presença será **OBRIGATÓRIA** para as Equipes que irão participar da Competição Baja SAE BRASIL 2024. Será registrada a **PRESENÇA** das equipes participantes em diferentes momentos durante a transmissão. As Equipes irão receber um convite por e-mail através da SAE BRASIL com as instruções para participação. Caso a Equipe não esteja na lista de e-mails da SAE BRASIL, por favor, encaminhar uma mensagem para: [thais.bozzo@saebrasil.org.br](mailto:thais.bozzo@saebrasil.org.br) ou entrem em contato pelo Telegram dos capitães (#fiquemespertos).

Atenciosamente,  
SAE BRASIL