











<b>PREDICTION TASK</b>  <p>Type of task? Entity on which predictions are made? Possible outcomes? Wait time before observation?</p> <p><i>Avaliação positiva ou negativa da estimativa de vendas mundial prevista para um videogame.</i></p> <p><i>A predição deve ocorrer antes de iniciar o processo de desenvolvimento do game.</i></p> <p><i>A observação do resultado real ocorrerá meses ou anos após a predição.</i></p>	<b>DECISIONS</b>  <p>How are predictions turned into proposed value for the end-user? Mention parameters of the process / application that does that.</p> <p><i>A predição nos dará a partir da quantidade total, global de vendas do videogame, a decisão se ele é ou não é considerado um sucesso de vendas. Consideramos um videogame vencedor no caso em que esta quantidade ultrapassa 1 milhão de cópias.</i></p>	<b>VALUE PROPOSITION</b>  <p>Who is the end-user? What are their objectives? How will they benefit from the ML system? Mention workflow/interfaces.</p> <p><i>Empreendedor que pretende lançar um videogame no mercado. Ele espera alcançar um nível considerável de vendas. Desta forma ele não investirá tempo de programação em um game que não lhe trará o retorno desejado.</i></p>	<b>DATA COLLECTION</b>  <p>Strategy for initial train set &amp; continuous update. Mention collection rate, holdout on production entities, cost/constraints to observe outcomes.</p> <p>Utilização de Datasets disponíveis na web com dados coletados no mercado de vendas de videogame.</p> <p>Rate do video_game_sales</p> <p>As informações contidas neste dataset são coletadas mundialmente o que é importante para qualquer iniciativa que seja direcionada para este mercado que tem como característica ser mundial.</p>	<b>DATA SOURCES</b>  <p>Where can we get (raw) information on entities and observed outcomes? Mention database tables, API methods, websites to scrape, etc.</p> <p><i>O formato dos dados foi disponibilizado em arquivos csv.</i></p> <p><a href="https://www.kaggle.com/datasets/gregorut/videogamesales">https://www.kaggle.com/datasets/gregorut/videogamesales</a></p> <p><a href="https://www.google.com/url?sa=t&amp;rct=j&amp;q=&amp;esrc=s&amp;source=web&amp;cd=&amp;ved=2ahUKEwiS84Dcicz4AhUhH7kGHVdGcnk4ChAWegQIEhAB&amp;url=https%3A%2F%2Fedisciplinas.usp.br%2Fpluginfile.php%2F5663800%2Fmod_resource%2Fcontent%2F1%2FVideoGameSales.pdf&amp;usq=AOvVaw00Vetwh2UEWLE5Mx07saB7">https://www.google.com/url?sa=t&amp;rct=j&amp;q=&amp;esrc=s&amp;source=web&amp;cd=&amp;ved=2ahUKEwiS84Dcicz4AhUhH7kGHVdGcnk4ChAWegQIEhAB&amp;url=https%3A%2F%2Fedisciplinas.usp.br%2Fpluginfile.php%2F5663800%2Fmod_resource%2Fcontent%2F1%2FVideoGameSales.pdf&amp;usq=AOvVaw00Vetwh2UEWLE5Mx07saB7</a></p>
<b>IMPACT SIMULATION</b>  <p>Can models be deployed? Which test data to assess performance? Cost/gain values for (in)correct decisions? <u>Fairness constraint</u>?</p> <p><i>Os modelos devem priorizar decisões mais precisas por ter um custo alto associado ao falso positivo.</i></p> <p><i>Em detrimento deveremos ter mais falsos negativos.</i></p>	<b>MAKING PREDICTIONS</b>  <p>When do we make real-time / batch pred.? Time available for this + featurization + post-processing? Compute target?</p> <p><i>Identificar como predição games que terão sucesso de vendas.</i></p>	<p><i>Para tomar a decisão de investimento, este empreendedor deverá coletar informações suficientes para permitir a avaliação do seu game a partir do modelo criado.</i></p> <p><i>De posse destas informações ele terá como retorno uma decisão binária: é um game vencedor, ou não é um game vencedor.</i></p>	<b>BUILDING MODELS</b>  <p>How many prod models are needed? When would we update? Time available for this (including featurization and analysis)?</p> <p><i>Model XBOSSGame</i></p>	<b>FEATURES</b>  <p>Input representations available at prediction time, extracted from raw data sources.</p> <p><i>Ano de lançamento do game</i></p> <p><i>Gênero do game</i></p> <p><i>Publicador do game</i></p> <p><i>Nota média de críticos</i></p> <p><i>Nota média de usuários</i></p> <p><i>Classificação ESRB</i></p>
	<b>MONITORING</b>  <p>Metrics to quantify value creation and measure the ML system's impact in production (on end-users and business)?</p> <p><i>Podemos ter como métrica uma análise da predição versus o resultado obtido globalmente para os casos considerados.</i></p> <p><i>Contagem de falsos positivos é a métrica mais importante que permite identificar o valor que a predição possui.</i></p>			

## Ready for the next step? Check out the ML Project Checklist!

Lead ML implementation with confidence with the CRISP—OWNML methodology (*Cross-Industry Standard Process to create your own Machine Learning system*) and its checklist. End-to-end ML projects are broken down into **9 phases of 4-5 tasks each**.

The checklist serves as a roadmap, listing in detail what you need to do, and in which order, so you can minimize risks and **make the most efficient use of your (and your team's) time**.

Learn more at [ownml.co/checklist](https://ownml.co/checklist)

