

# Machine Learning Canvas (Big Data MBA Version)

Designed by: Felipe Cunha

Date: 06/05

Iteration:

V2.0

<b>Decisions</b> How are predictions used to make decisions that provide the proposed value to the end-user?  Prever o perfil de gasto energético da indústria nos meses seguintes visando obter um melhor aproveitamento da energia e consequente diminuição dos custos de produção e das emissões de CO <sub>2</sub> ;	<b>ML task</b> Input, output to predict, type of problem.  <b>Inputs:</b> corrente, fator de potência, emissão CO <sub>2</sub> , data, dia da semana... <b>Outputs:</b> consumo de energia elétrica no mês <b>Type:</b> regressão	<b>Value Propositions</b> What are we trying to do for the end-user(s) of the predictive system? What objectives are we serving?  Fornecer subsídios para melhor controlar o uso de energia elétrica na indústria, visto que esse é um fator crítico tanto para a produção (redução de custos, otimização da utilização de equipamentos, planejamento para a produção, etc). Além disso, um gasto mais eficiente de energia elétrica é imprescindível para a proteção do meio ambiente;	<b>Data Sources</b> Which raw data sources can we use (internal and external)?  Internal: dados de uma indústria nos últimos 12 meses com medidas de 15 em 15 minutos; External: a princípio, não será necessário	<b>Collecting Data</b> How do we get new data to learn from (inputs and outputs)?  A princípio não será necessário.
<b>Making Predictions</b> When do we make predictions on new inputs? How long do we have to featurize a new input and make a prediction?  A resposta do modelo será avaliada mensalmente; O modelo deverá ser retreinado idealmente mês a mês, com prazo máximo de 2 em 2 meses.	<b>Offline Evaluation</b> Methods and metrics to evaluate the system before deployment.  Serão utilizadas métricas relacionadas à modelos de regressão: coeficiente de determinação e raiz do erro quadrático médio; Também se levará em conta o erro absoluto (mês a mês);		<b>Features</b> Input representations extracted from raw data sources.  Variáveis relacionadas ao consumo energético da indústria (correntes, fatores de potência, emissões de CO <sub>2</sub> , dia da semana) ao longo de 1 ano com medidas realizadas de 15 em 15 minutos;	<b>Building Models</b> When do we create/update models with new training data? How long do we have to featurize training inputs and create a model?  O modelo deverá ser retreinado a cada mês, ou no máximo com 2 meses;
<b>Prescription</b> Once we have a prediction, what do we do?	<b>Automation</b> How to we automate standard procedures with the prescriptive insights?		<b>Live Evaluation and Monitoring</b> Methods and metrics to evaluate the system after deployment, and to quantify value creation.  A principal métrica a ser avaliada deve ser o erro absoluto mês a mês aliado ao RMSEP	