**Projeto Aplicado - Proposta de Solução**

|  |  |
| --- | --- |
| **Nome do Aluno** | **ADRIANO MARABUCO DE ALBUQUERQUE LIMA** |
| **Curso** | **MBA EM APRENDIZADO DE MÁQUINA** |
| **Linha de Especialização** | **SISTEMAS DE RECOMENDAÇÃO** |
| **Orientador** | **BRUNO RAFAEL DE OLIVEIRA RODRIGUES** |
| **Data** | **24/04/19** |

**EXPLORANDO A OPORTUNIDADE**

1. Apresente a oportunidade vislumbrada que motivou o desenvolvimento da solução:
   1. Quais são os principais desafios (“dores”) que o projeto pretende solucionar?
   2. Quais são as ações dos clientes que evidenciam os desafios?
   3. Quais hipóteses sobre o desafio foram validadas? Como foi feita a validação?
   4. Anexe a(s) ferramenta(s) pertinente(s) utilizadas nesta etapa: matriz CSD, Business Design Blueprint, Experiment Board, entre outras.

A gestão de investimentos lida com diversos desafios cujos impactos podem levar o negócio a ruína. Riquezas acumuladas durante muitos anos podem sumir em questão de dias. Os gestores de investimentos vivem na “corda bamba” entre ganhar fortunas e perdê-las pelos motivos que se seguem: alterações súbitas no comportamento de um ativo; exposição superestimada ou subestimada do risco; e falhas no momento de compra ou venda.

Os investimentos de renda variável são caracterizados pela volatilidade dos preços dos ativos, o que é resultado das variações de preços das compras e vendas. O problema é que essa variável também oscila com o tempo, portanto tem uma distribuição não estocástica, gerando grande incerteza nas alterações dos preços, pois eles podem sair de um patamar de baixa volatilidade para alta de um dia para o outro, e um investimento de baixo de risco tornar-se de alto. Essas alterações súbitas das variações de preços podem tornar um investimento rentável em um fracasso com grandes perdas financeiras.

Essa ruptura no padrão de negociação, e risco, gera outro problema na gestão dos investimentos, pois interfere no binômio risco/retorno do ativo, e de forma agregada no portfólio de ativos, criando uma exposição indesejada ao risco. Investidores agressivos buscam maiores retorno e estão dispostos a suportarem mais riscos, enquanto o oposto vale para os conservadores. Ao definir a estratégia de investimento, o gestor busca adequar os investimentos ao perfil do público-alvo dos investidores, entretanto mudanças no mercado podem provocar um desalinhamento disso. Por exemplo, o gestor pode comprar determinada ação buscando montar uma carteira de baixo risco e baixo retorno para investidores mais conservadores, e por um aumento repentino na volatilidade sofrer grandes perdas monetárias com a queda no preço dos ativos que se tornarem instáveis.

Outra questão que também deriva da volatilidade é o timing de compra e venda dos ativos. A fórmula para o sucesso no mercado financeiro é simples: comprar na baixa e vender na alta. O problema é que na prática é muito difícil acertar a hora de comprar e vender, e muitas vezes acontece o oposto. Após um alta contínua nos preços, muitos negociadores resolvem comprar o ativo achando que a alta vai persistir, e acaba acontecendo o contrário, ou após uma grande queda, muitos ficam receio de a queda continuar e se desfazem de suas ações, ficando com o prejuízo e perdendo a recuperação dos preços. Essas falhas de timming dilapidam o patrimônio dos fundos de investimento.

Os desafios elencados anteriormente são evidenciados por casos notáveis de crises em fundos de investimento, assim como pelas informações de mercado.

No que tange a variação do padrão de comportamento dos ativos e seus efeitos é ilustrativo o caso do LTCM, que foi um grande fundo de investimento americano idealizado e gerenciado por economistas ganhadores do prêmio do Nobel (Myron S. Scholes e Robert C. Merton), que usavam estratégias de investimento baseadas em modelos econométricos avançados, e foram surpreendidos por mudanças no comportamento do mercado que provocaram um caos financeiro e a bancarrota do fundo e muitos investidores:

*Initially successful with annualized return of over 21% (after fees) in its first year, 43% in the second year and 41% in the third year, in 1998 it lost $4.6 billion in less than four months following the 1997 Asian financial crisis and 1998 Russian financial crisis, requiring financial intervention by the Federal Reserve, with the fund liquidating and dissolving in early 2000.  
Fonte: https://en.wikipedia.org/wiki/Long-Term\_Capital\_Management.*

A alteração no binômio risco/retorno pelas mudanças no comportamento dos ativos pode ser identificada pela análise de alguns casos, como o da Petrobrás, conforme o gráfico a seguir.

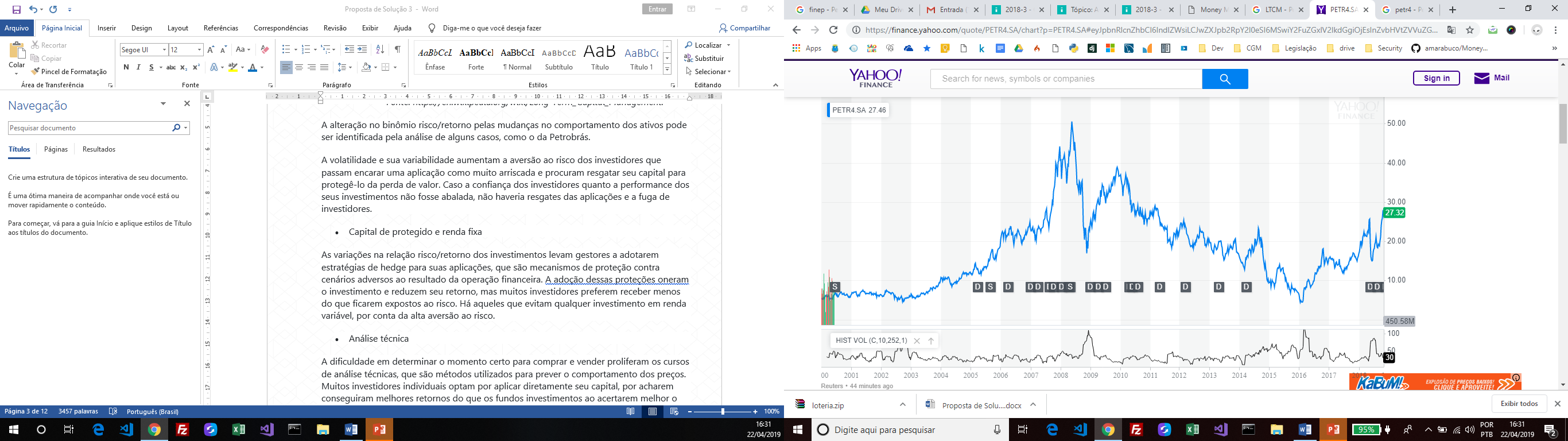
**Preço**

**Volatilidade**

**2**

**1**

Figura 1 Preço e volatilidade PETR4 - 2000 a 2019



No primeiro momento observa-se que o preço do ativo mantém uma alta consistente com baixa volatidade, revelando um perfil de baixo risco, enquanto no segundo momento nota-se baixas e altas alternadas com alta volatilidade, mostrando que o perfil mudou para alto risco, quando comparado ao primeiro momento. Como mencionado anteriormente, a relação risco/retorno do ativo mudou ao longo dos anos, então o gestor de investimentos que comprou a estável no Petrobrás nos anos de 2003 a 2007 foi surpreendido por um pico de volatilidade em 2009, e um aumento consistente na volatilidade a partir de 2014, tornando-o um ativo instável.

A evidência do timming também pode ser inferida do gráfico exposto. Quem comprou no final de 2007 e vendeu em meados de 2008, auferiu um bom lucro, porém quem esperou para vender em 2009 pode ter saído no prejuízo, e quem se empolgou com a alta dos preços e comprou em meados de 2008, depois se assustou com a queda e vendeu no início de 2009 ficou um grande prejuízo, mas se esperou para vender em 2010 reduziu sua perda.

Diante do exposto, percebe-se que as alterações súbitas no comportamento de um ativo; exposição superestimada ou subestimada do risco; e falhas no momento de compra ou venda são dores reais sofridas pelos investidores do mercado financeiro.

A utilização do aprendizado de máquina para identificar pelos dados históricos: os padrões de comportamento dos ativos, antecipar suas variações e prever preços pode oferecer uma forma de mitigar os riscos das decisões de investimentos e facilitar a vida dos gestores e investidores.

Existem empresas já trabalhando com esse tipo de solução, conforme relata a matéria do jornal valor econômico “Inteligência artificial facilita a vida de investidores” de 31/01/19:

“Seival Investimentos e Visia estão entre as gestoras de fundos baseados na tecnologia. Em 2018, o Seival FGS, por exemplo, rendeu 41% com modelos quantitativos seguidores de tendências, comemora o co-fundador Carlos Chaves. A Visia, com patrimônio sob gestão perto de R$ 900 milhões, busca ineficiências de mercado em repetição de padrões e usa aprendizado de máquina (machine learning) para executar ordens de compra sem impactar o mercado, explica o sócio Flavio Terni. Seu fundo Zarathustra rendeu 240% do Certificado de Depósito Interfinanceiro (CDI) em 2018.

A falta de familiaridade com conceitos e lógica das finanças colabora para a popularização dos robôs para aconselhamento. Para o CEO da Warren, Tito Gusmão, 85% dos investidores brasileiros preferem delegar decisões de investimentos por falta de conhecimentos suficientes sobre o tema.

A fintech, nascida em 2017, já atraiu mais de 50 mil clientes, maneja perto de R$ 300 milhões e usa a tecnologia para identificação de perfil e objetivos de investimento, recomendação e back office. Oferece quatro fundos próprios e, a partir de fevereiro, fundos de terceiros, graças à aquisição de uma corretora com recursos do aporte de um fundo do Vale do Silício. O serviço de gestão do portfólio de investimentos do cliente custa 0,5% ao ano, e produtos próprios têm taxa zero.

...

Já os serviços de gestão de patrimônio da Magnetis são sustentados por plataforma digital automatizada para entender o cliente, recomendar investimentos, definir carteiras e, se necessário, modificá-las. "A automatização cria eficiência operacional e deixa o serviço mais acessível", justifica o CEO Luciano Tavares. São 0,4% ao ano pela gestão e taxa administrativa de 0,3% a 0,5% para alocação nos fundos, com atendimento humano para os investidores.

A fintech Vérios, por sua vez, cobra entre 0,4% e 0,65% ao ano sobre o valor investido - somando custos como corretagem, custódia e outros, o valor total fica em 0,95% anuais. Robôs ajudam no gerenciamento da conta individual e no cálculo da alocação, com definição final pelo time da casa. No ano passado, o rendimento ficou entre 105% e 135% do CDI, detalha o CEO Felipe Sotto-Maior.”

Fora esse artigo nota-se por pesquisas na web que é muito grande a quantidade de projetos e artigos que tratam do aprendizado de máquina aplicado ao previsão de séries temporais de investimentos (ações, derivativos, criptomoedas).

Assim, fica comprovado que a utilização de sistemas de recomendação para investimentos é uma tendência forte no mercado financeiro, e um fundo de investimento gerenciado por algoritmos seria capaz de atrair muitos investidores.

A matriz CSD  permite  explorar  três  questões  essenciais  sobre  seu  projeto:  i)  as certezas, ou seja, o que já se sabe sobre o problema/desafio; ii) as suposições, ou seja, as hipóteses relacionadas ao problema/desafio; e iii) as dúvidas, de forma que ainda existem perguntas a serem respondidas.

Nessa matriz também é possível analisar o problema sob diferentes óticas, sendo as mais  comuns:  i)  atores,  que  são  as  pessoas  envolvidas  no  problema;  ii) cenários, onde ocorrem os problemas; e iii) regras, que representam a relação entre os atores e os cenários.

Tabela 1 Matriz CSD

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Certezas** | **Suposições** | **Dúvidas** |
| **Investidores e gestores de investimentos** | * Alta demanda por maior retorno para investimento | * Demanda por inovações no setor de fundo de investimentos | * Aceitação em delegar investimento para gestão por máquinas |
| **Aplicação de IA ao mercado financeiro** | * Avanços tecnológicos mudaram o mercado financeiro | * Capacidade dos algoritmos de prever os movimentos do mercado | * Geração de resultados consistentes de investimentos efetuados por robôs |
| **Regras** | * Utilização da tecnologia para apoiar o trabalho dos gestores de fundos | * Uso crescente da Inteligência Artificial na gestão de capital | * Fundos de investimento baseados em tecnologia vão dominar o setor financeiro |

1. Apresente os benefícios da solução que motivarão a compra de seu produto/serviço ou a adoção dentro de sua organização.
   1. Quais são os principais benefícios da solução (redução de custos, aumento de receita, ganho de imagem, redução de impactos ambientais, etc.)?
   2. Anexe a(s) ferramenta (s) pertinente (s): Business ModelCanvas, Canvas de Proposta de Valor, Matriz de concorrência, entre outras.

Um fundo gerenciado com base em sistemas de recomendação de machine learning apresenta diferenciais relevantes em relação aos tradicionais fundos de investimento.

Tabela 2 Fundo Tradicional x Fundo com Inteligência Artificial

|  |  |
| --- | --- |
| Fundo tradicional | Fundo com inteligência artificial |
| Composto por analistas de investimentos e negociadores. Os analistas avaliam o mercado para e gerar recomendações para suportar a tomada de decisão de investimento, enquanto os negociadores executam as operações de compra e venda. | Composto por cientistas de dados e robôs de análise e negociação. Os cientistas de dados preparam os dados e desenvolvem os robôs para efetuarem as análises e negociações. |
| • Analistas de investimento podem demorar a emitir relatório sobre investimento ao lidarem com grandes massas de informação e eventos simultâneos e/ou contraditórios. | • Análises complexas e rápidas com uso intensivo de dados e poder computacional oferecem menor tempo de resposta em relação às equipes de análise. |
| • A recomendação de investimento é influenciada por fatores emocionais e vieses dos gestores e negociadores | • A recomendação de investimento é baseada em cálculos e probabilidades sem interferência emocional ou vieses pessoais |
| • Negociadores contam com intuição e experiência para acertar os melhores momentos para comprar e vender | • As negociações são tomadas com base em dados e estatísticas |
| • Transações são executadas pela digitação de dados pelos negociadores na plataforma de negociação. | • Sistemas automatizados comunicam-se com mais velocidade com a plataforma e não estão sujeitos a erro de digitação. |

O resultado desses diferenciais seriam melhor performance dos investimentos e mais retorno para os investidores. Além disso, o custo da equipe do fundo tende a ser menor com a substituição de negociadores e analistas por robôs, pois profissionais experientes e com histórico de bons resultados exigem alta remuneração, onerando a folha de pagamento do fundo de investimento.

O Business Model Canvas (BMC), ou simplesmente CANVAS, proposto por Osterwalder e Pigneur (2010), é uma ferramenta de fácil visualização e compreensão, que auxilia na construção, evolução e direcionamento do modelo de negócio de um empreendimento.

O método divide o negócio em nove blocos principais, sendo que cada bloco corresponde a um ponto chave do empreendimento, permitindo a obtenção de uma visão global do negócio e, consequentemente, incrementando a capacidade de controle sobre ele.

Tabela 3 Business Model Canvas

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parcerias Principais   * Corretoras * Bancos * Family offices | Atividades Principais   * Gestão de investimentos * Desenvolvimento de algoritmos para mercado financeiro | Proposta de valor   * Utilização de alta tecnologia para gestão de recursos próprios e de terceiros | Relacionamento com clientes   * Atendimento personalizado * Atendimento online | Segmentos de clientes   * Investidores institucionais * investidores pessoa física |
| Recursos principais   * Sistemas de inteligência artificial * Relatório analíticos * Equipe de analistas de negócios * Equipe de desenvolvedores | Canais   * Site * Equipe de vendas |
| Estrutura de custos   * Salários * Taxas (corretagem, custódia) * Infraestrutura (escritório, computadores, servidores) * Serviços (contabilidade, jurídico, TI) | | | Fontes de receita   * Taxa de administração * Taxa de performance | |

**SOBRE A SOLUÇÃO**

1. Construção do protótipo/MVP e novas versões:
   1. Quais são os requisitos básicos para construção do MVP?
   2. Qual é o plano de ação para desenvolvimento do MVP?
   3. Como o MVP será validado? Descreva o experimento e métricas de validação.
   4. Qual é a evolução pretendida para a solução (TRM)?

A construção de um protótipo/MVP permite tornar a oportunidade mais viável e tangível, de forma que constitui um passo crítico na criação do negócio. Por meio dele é possível ter um “embrião” do serviço, testar seu funcionamento e cumprimento de seus objetivos.

Os requisitos para o desenvolvimento do MVP seriam: obter dados do mercado de renda variável, algoritmos de aprendizado de máquina e infraestrutura computacional.

Os dados do mercado (dataset) incluem a variação dos preços nos pregões diários e informações econômico-financeiras que podem ser obtidas de diversas fontes: bovespa, yahoo, cvm, google. Tais fontes permitem fazer o download de arquivos csv. As informações serão utilizadas para o treinamento dos algoritmos de predição. O fornecimento de dados via API pela fonte é um atributo desejável, pois facilita o fluxo de informações, tornando a aplicação mais flexível e responsiva.

Os algoritmos são os modelos computacionais utilizados para a aprendizagem de máquina fornecidos por diversas bibliotecas abertas de machine learning, como scikitlearn, keras, tensorflow que compõem o ecossistema de Data Science do Python. Tais modelos devem ser implementados pela leitura e análise do dataset para gerar as funções de aproximação que serão usados na previsão dos preços. O treinamento consiste nessa etapa de implementação em que se é necessária uma infraestrutura de computadores para executar os métodos dos algoritmos.

A infraestrutura refere-se ao poder computacional necessário para executar o processamento dos dados pelos algoritmos, de forma a gerar o aprendizado. Quanto maior o volume de dados e a complexidade do algoritmo, maior necessidade de poder computacional (memória e processamento). Se forem utilizadas análises em tempo real, o que aumenta a velocidade dos dados, também é necessário mais poder. A utilização da computação em nuvem permite trabalhar com um custo e poder computacional variável, sendo ideal para o ambiente de desenvolvimento em que o MVP se insere. Algumas plataformas de cloud computing como, a Heroku, possuem pacotes de serviços gratuitos para desenvolvimento e teste da aplicação, além de garantir a escalabilidade para quando a aplicação estiver em produção.

A construção do MVP começa pelo planejamento das ações necessárias para o desenvolvimento da aplicação. Depois disso, o plano de ação fornecerá a orientação devida para avançar na execução do projeto.

1. Levantamento dos dados
   1. Yahoo: download CSV
   2. Google Finance + Pandas data-reader: via API
   3. Histórico de negociação dos últimos 5 anos ou mais
   4. Ações com alta liquidez que compõem o IBOVESPA
2. Seleção de modelos de aprendizado de máquina
   1. Pré-processamento dos dados
   2. Regression Decision Tree
   3. Random Forest
   4. Support Vector Machines
   5. Multi-layer Perceptron Regressor
   6. Long Short-Term Memory Neural Network
3. Treinamento dos modelos e teste de previsão
   1. Definição de parâmetros dos modelos
   2. Execução do treinamento
   3. Previsão e avaliação dos resultados
4. Classificação dos melhores modelos
   1. Ranking de acurácia dos modelos
   2. Seleção dos modelos mais eficazes

A validação do MVP deve ser feita por meio dos testes e análise do seu poder preditivo, quanto mais acurada a previsão de preço, mais o MVP atinge seu objetivo, que é provar ser capaz de prever o comportamento do mercado, e, portanto, ser capaz de recomendar investimentos.

Os testes podem ser efetuados em dois horizontes temporais. O primeiro seria o backtest que utiliza os dados históricos para fazer a previsão, e assim conferir se as previsões estariam corretas. Esse teste mostra o risco do underfitting, que significa que o modelo não foi capaz de prever corretamente os preços. O segundo seria testar com "dados futuros", ou seja, a medida que novos dados surgirem eles seriam usados para testar a previsão do modelo. Esse teste mostra o risco do overfitting, que significa que o modelo se especializou tanto nos dados de aprendizagem que não é capaz prever corretamente quando utiliza dados novos e diferentes, portanto não é útil.

Caso o MVP passe os testes simulados, uma outra forma de validação seria utilizá-lo em investimentos efetivos, o que demonstraria sua capacidade gerar resultados financeiros reais, e não somente simulados. Nesse aspecto algo que pode acontecer é que se os investimentos forem vultuosos, eles modem interferir na capacidade preditiva do modelo, pois poderiam gerar um novo padrão de comportamento que não havia sido incorporado ao modelo.

É importante ressaltar que o treinamento do modelo deve ser contínuo, mesmo que ele seja aprovado na validação, pois como as condições de mercado mudam e o modelo deve estar sempre se adaptando a elas, sob pena de perder seu poder preditivo.

Algumas métricas que podem ser usadas na validação do MVP são:

* Mean Absolute Error

A função mean\_absolute\_error calcula o erro absoluto médio, uma métrica de risco correspondente ao valor esperado da perda de erro absoluta ou perda de norma.

Se é o valor previsto da i-ésima amostra, e é o valor verdadeiro correspondente, então o erro absoluto médio (MAE) estimado é definido como

* Mean Squared Error:

A função mean\_squared\_error calcula o erro quadrático médio, uma métrica de risco correspondente ao valor esperado do erro ou perda quadrática (quadrática).

Se é o valor previsto da i-ésima amostra, e é o valor verdadeiro correspondente, então o erro quadrático médio estimado (MSE) é definido como

O MSE pode ser usado com Gradient Boosting regression para aprimorar a qualidade do modelo.

* R² score

R², o coeficiente de determinação fornece uma medida de quão bem as amostras futuras provavelmente serão previstas pelo modelo. A melhor pontuação possível é 1.0 e pode ser negativa (porque o modelo pode ser arbitrariamente pior).

Se i é o valor previsto da i-ésima amostra, e é o valor verdadeiro, então o R² é definido como:

O roadmap do negócio representa o planejamento da evolução do negócio ou aplicação. No caso do fundo de investimento, ele é determinado pelo aumento progressivo da complexidade e atividades executadas pelos algoritmos.

O primeiro passo é desenvolver o sistema de previsão de preços, e a progressão é rumo a realização das operações de compra e venda automatizadas criando um sistema completo para a gestão do fundo de investimento.

No caso do modelo preditivo, ele acompanharia a evolução do negócio incorporando mais dados, atividades ou algoritmos, de acordo com os testes e avaliações a serem feitos, os quais se pode antecipar sem a conclusão do MVP.

1. Como vender seu projeto:
   1. Quais são os principais indicadores econômico-financeiros do projeto?
   2. Apresente o Círculo Dourado e a Matriz de Stakeholders da sua solução.

Os indicadores econômico-financeiros são essenciais para avaliar a viabilidade de um projeto, de forma que se possam prever e monitorar seus custos e receitas, e evitar um projeto cujo custo seria superior ao valor gerado, configurando um desperdício de recursos (tempo, capital e pessoas).

A métrica base de avaliação do projeto seria o retorno sobre investimento (ROI), por meio dela seria capaz aferir a qualidade dos resultados obtidos com aplicação de aprendizado de máquina. Esse indicador pode ser utilizado tanto para o projeto e investimentos realizados com auxílio do modelo preditivo. Quanto maior o ROI do projeto maior a atratividade para seus sócios, e quanto aos investimentos, melhor para os clientes.

Outra métrica a ser utilizada é valor presente líquido (VPL), pois a avaliação somente em relação ao ROI não é suficiente já que ela é pode gerar um ROI alto em baixa escala e um ROI baixo em alto escala, e assim não demonstrar o valor monetário gerado como um todo. Por isso seria necessário utilizar também o VPL para aferir o valor presente do projeto e quantificar o seu ganho ou prejuízo dentro do prazo considerado. Se o ROI diminuir com o aumento da escala do projeto, mas o VPL ainda for positivo, o projeto mantém-se viável.

Não há como estimar o ROI e VPL antes da conclusão do MVP, pois sem ele não há perspectiva de receita nem gastos operacionais, por outro lado o custo de elaboração do MVP é considerado desprezível para o fluxo de caixa do negócio.

A metodologia do círculo de ouro torna possível desenvolver o valor de uma nova ideia ou negócio, sistematizando um novo jeito de pensar, agir e comunicar de maneira mais eficaz.

O mapa de Stakeholders trata-se de uma ferramenta utilizada na abordagem do Design Thinking para mapear as pessoas que interferem direta ou indiretamente no projeto.

|  |  |
| --- | --- |
| **Manter satisfeito**   1. Sócios (a serem captados) 2. Funcionários (a serem contratados) 3. Investidores | **Gerenciar de perto**   1. Comissão de Valores Mobiliários |
| **Monitorar**   1. Concorrentes (Seival Investimentos e Visia) | **Manter informado**   1. Corretoras 2. Bancos 3. Receita Federal |