**Aula 02 - 01/04/2025 - Fundamentos de Machine Learning**

Resumo

A aula ministrada por Cristiano Ritta e Jonas Tomazi abordou fundamentos de Machine Learning, incluindo tipos de aprendizado, reconhecimento de padrões, e a diferença entre programação tradicional e IA, além da apresentação do projeto de IA de Marco Aurélio. Foram demonstrados exemplos práticos com Google Colab, APIs da OpenAI e Serper, utilizando agentes autônomos para busca e geração de relatórios, enfatizando a importância da identificação de padrões e o uso de ferramentas como Obsidian. Como próximos passos, os participantes deverão configurar suas chaves API e explorar os agentes autônomos em projetos futuros.

Detalhes

* **Apresentação de Marco Aurélio e sua inovação em IA:** Jonas Tomazi e Cristiano Ritta apresentaram Marco Aurélio, um aluno que desenvolveu um GPT conectado a uma base de dados de 400 livros, gerando respostas precisas. Eles destacaram a inovação do projeto, considerando-o algo fantástico e revolucionário no contexto das LLMs.
* **Recursos da pós-graduação em IA da Data Virtus:** A Data Virtus oferece aos alunos suporte para a resolução de problemas e criação de soluções em IA, com o objetivo de promover a colaboração entre eles (00:01:12). Eles fornecem materiais complementares como anotações em Obsidian (em formato markdown), slides, gravações e mapas mentais (00:02:32).
* **Plugin de mapa mental para Obsidian:** Cristiano Ritta mencionou um plugin do Obsidian para geração de mapas mentais, chamado "Enging mind map". Jonas Tomazi demonstrou interesse em aprender mais sobre o plugin (00:04:37).
* **Tópicos da aula sobre fundamentos de Machine Learning:** A aula abordou os fundamentos de Machine Learning, incluindo a relação entre IA e Machine Learning, o processo de aprendizado de máquina, tipos de aprendizado (supervisionado, não supervisionado, por reforço), pré-processamento de dados e treinamento de IA (00:05:23).
* **Aprendizado de máquina como reconhecimento de padrões:** Cristiano Ritta explicou o aprendizado de máquina como um subcampo da IA relacionado ao reconhecimento de padrões, comparando-o ao aprendizado humano (00:08:07). Eles enfatizaram que, mesmo em aprendizado profundo e redes neurais, o reconhecimento de padrões é fundamental (00:09:25).
* **Conceitos-chave de Machine Learning:** Foram discutidos três elementos essenciais do Machine Learning: experiência, tarefa e desempenho (00:13:17). Um exemplo citado foi o treinamento de um filtro de spam com milhares de e-mails para identificar padrões e melhorar a precisão (00:14:41).
* **Prêmio Nobel e a relação entre cérebro humano e IA:** O prêmio Nobel de física para John Hopfield e Geoffrey Hinton, pais da rede neural artificial, foi mencionado como reconhecimento da base teórica para o aprendizado em computadores, inspirada no cérebro humano (00:17:10).
* **Diferenças entre programação tradicional e IA:** A programação tradicional se baseia em dados e regras explícitas, enquanto a IA aprende padrões nos dados sem regras pré-definidas (00:18:43). Cristiano Ritta utilizou exemplos com planilhas para ilustrar a diferença (00:21:15).
* **Exemplo do robô da Tesla e reconhecimento de padrões:** Jonas Tomazi compartilhou um exemplo de um robô da Tesla que aprendeu a ganhar no jogo pedra-papel-tesoura, identificando padrões de comportamento humano (00:24:58).
* **Inteligência artificial e reconhecimento de padrões:** Eles concluíram que a IA se baseia principalmente no reconhecimento de padrões, aplicável em diversas áreas, como previsão do tempo (00:25:58). Cristiano Ritta destacou a importância da identificação de padrões e a ausência de programação explícita como elementos cruciais da IA (00:27:35).
* **Tipos de aprendizado de máquina:** A aula cobriu aprendizado supervisionado (dados rotulados), não supervisionado (dados sem rótulos) e por reforço (feedback do usuário) (00:34:42) (00:41:20). A importância dos dados rotulados para o aprendizado supervisionado foi enfatizada (00:36:26). Eles discutiram também a vantagem do aprendizado semi-supervisionado para reduzir custos e aumentar eficiência (00:44:11).
* **A influência da língua inglesa nos modelos de IA:** A predominância do inglês nos dados de treinamento afeta a performance dos modelos em outras línguas, sendo o inglês geralmente o idioma de melhor performance (00:47:19). Cristiano Ritta confirmou essa observação e deu o exemplo do recurso "thinking" do ChatGPT que opera em inglês (00:48:02).
* **Aprendizado por transferência e Fine-tuning:** A aula mencionou o aprendizado por transferência, onde o conhecimento de um modelo é aplicado a outros, e o fine-tuning, um ajuste fino de um modelo pré-treinado (00:48:50).
* **Declaração de Elon Musk sobre o modelo Grock:** Cristiano Ritta comentou sobre a declaração de Elon Musk a respeito do modelo Grock da xAI, que teria consumido todo o conhecimento humano dos últimos 4000 anos e a aquisição do Twitter (X) pela empresa como consequência dessa capacidade (00:51:07).
* **Atualização de conhecimento em modelos de IA:** Eles discutiram o desafio de manter os modelos de IA atualizados com o conhecimento humano em constante evolução, mencionando a utilização de busca na internet em tempo real como solução (00:53:10).
* **Transformers e o mecanismo de atenção:** A aula introduziu os modelos Transformers e seu mecanismo de atenção, que permite focar nos elementos mais relevantes de um conjunto de dados, impactando a qualidade das respostas (00:54:05).
* **GPTs e a geração de texto:** Foi explicado como os GPTs geram texto através do cálculo probabilístico da próxima palavra em uma sequência, baseado em padrões identificados (00:56:50). A combinação de modelos para gerar diferentes tipos de conteúdo (texto, imagem, áudio, etc.) também foi abordada (00:58:03).
* **Tokenização e Identificação de Padrões:** Cristiano Ritta explicou o sistema de tokenização e sua aplicação em diversas áreas, como processamento de linguagem natural, reconhecimento de imagem e áudio. Eles destacaram a importância da identificação de padrões para a aplicação prática em segurança pública e análise de previsão de crimes (00:59:15). A aplicação prática incluirá reconhecimento de voz, facial e de objetos, demonstrando a utilidade do *machine learning* (01:00:15).
* **Reações Positivas à Aula:** Jonas Tomazi expressou entusiasmo pela aula, descrevendo-a como "fantástica" e "top" (01:00:15), elogiando o conhecimento de Cristiano Ritta e a integração do conteúdo no Obsidian. Eles anteciparam a expansão do "segundo cérebro" com o acúmulo de informações (01:00:58).
* **Prática com Google Colab e APIs:** Jonas Tomazi guiou os participantes em um exercício prático usando o Google Colab, solicitando que eles acessassem um link e salvassem uma cópia no *drive* (01:02:40). Eles enfatizaram a importância de salvar uma cópia para edição (01:03:39).
* **Integração de Tutorial e Código:** Foi explicado que o tutorial e o código foram integrados para facilitar a reprodução da aula pelos participantes. O objetivo do exercício era simular uma busca na internet e gerar um relatório usando agentes de IA (01:04:38).
* **Acesso à OpenAI e Serper APIs:** Jonas Tomazi instruiu os participantes a criar uma conta na OpenAI e acessar a
* a perspectiva de automatizar a geração de código e times de agentes de IA (01:43:13).
* **Sorteio do Livro:** Ao final, Jonas Tomazi realizou um sorteio do livro do Cristiano Ritta entre os participantes que comentaram no vídeo do YouTube . Os ganhadores foram Erman e Moisés (01:49:15).

**Resolução de Dúvidas e Considerações Finais:**  Foram respondidas perguntas dos participantes, incluindo sobre configurações de computador para IA e a possibilidade de usar os agentes em diferentes plataformas .  Jonas Tomazi reforçou que o objetivo era quebrar o gelo e apresentar conceitos iniciais.