Astronomia na era do Big Data

Vocês sabiam que o James Webb Space Telescope gera cerca de 235 GB de dados por dia?

2026-03-02



Figure 1: Capa do artigo

Vocês sabiam que o James Webb Space Telescope gera cerca de 235 GB de dados por dia?

E que o Vera C. Rubin Observatory deve produzir algo em torno de 36 terabytes por noite em 2026?

Astronomia foi uma das primeiras disciplinas a estar definitivamente na era do Big Data.

Big Data é definido pelos 3 Vs:

Volume: quantidade massiva de dados.

Velocidade: dados sendo gerados continuamente.

Variedade: imagens, espectros, medições de brilho, posição e séries temporais.

Os telescópios modernos operam exatamente nesse cenário.

E sabe o que é humanamente impossível? analisar tudo que está no céu.

Nenhum grupo de cientistas consegue analisar manualmente milhões de imagens do céu.

Por isso aprendizado de máquina virou parte da infraestrutura científica.

© Artemisia Weyl 2024–2026 — Todos os direitos reservados.

Hoje ele é ferramenta central na astronomia.

Existem dois tipos principais de aprendizado de máquina usados na astronomia:

- Aprendizado supervisionado
- Aprendizado não supervisionado

No aprendizado supervisionado, o modelo aprende com exemplos já rotulados por especialistas.

Tipo:

Imagem de galáxia já classificada.

Curva de luz já identificada como exoplaneta.

Espectro já associado a uma distância específica.

O modelo aprende essa relação entre entrada e resposta esperada.

No aprendizado não supervisionado, não existem rótulos prévios.

O computador organiza os dados por similaridade.

Ele agrupa objetos parecidos e pode revelar padrões que ainda não tinham sido identificados pelos pesquisadores.

É muito usado para explorar grandes bases e encontrar estruturas escondidas.

Mesmo com os dois sendo importantes, o supervisionado virou peça central.

Porque muitas perguntas astronômicas já têm um alvo definido:

Qual a estrutura da galáxia?

Qual o desvio pro vermelho?

Há lente gravitacional nessa imagem?

Existe sinal de planeta nessa curva de luz?

Nesses casos, o modelo aprende a prever exatamente essas perguntas definidas.

Hoje redes neurais convolucionais analisam diretamente os pixels das imagens.

Transformers e outras arquiteturas modernas alcançam níveis de precisão superiores a 90% em tarefas de classificação astronômicas.

Sem esses modelos, analisar o volume gerado por telescópios como o James Webb seria impraticável.

Gostaram dessa curiosidade? Comenta pra gente

© Artemisia Weyl 2024–2026 — Todos os direitos reservados.