

1 Projetos

1. Classificação de roupas — Fashion-MNIST

- **Tarefa:** classificar 10 categorias de roupas.
- **Dataset:** Fashion-MNIST
- **ML clássico:** SVM, Random Forest.
- **Extensão DL:** CNN, ResNet simplificada.
- **Extensão RL:** atenção seletiva, bandit para aumento de dados.

2. Classificação de imagens naturais — CIFAR-10/100

- **Tarefa:** classificar 10 ou 100 classes de objetos e animais.
- **Dataset:** CIFAR-10/CIFAR-100
- **ML clássico:** HOG + SVM, Random Forest.
- **Extensão DL:** CNN, ResNet, EfficientNet, Vision Transformers (ViT).
- **Extensão RL:** controle de drone, NAS.

3. Filtro de spam — SMS Spam Collection

- **Tarefa:** classificar mensagens como spam ou não spam.
- **Dataset:** SMS Spam Collection
- **ML clássico:** Naive Bayes, Regressão Logística, SVM.
- **Extensão DL:** LSTM, GRU, BERT.
- **Extensão RL:** bandit contextual.

4. Análise de sentimento — IMDB

- **Tarefa:** classificar críticas como positivas ou negativas.
- **Dataset:** IMDB
- **ML clássico:** TF-IDF + Regressão Logística, SVM.
- **Extensão DL:** CNN para texto, LSTM, Transformers (BERT).
- **Extensão RL:** RLHF, agente de diálogo.

5. Classificação de tópicos de notícias — 20 Newsgroups

- **Tarefa:** classificar notícias em categorias.
- **Dataset:** 20 Newsgroups
- **ML clássico:** Naive Bayes, SVM.
- **Extensão DL:** BERT.
- **Extensão RL:** recomendação de notícias.

6. Detecção de fraude — Credit Card Fraud Detection

- **Tarefa:** classificar transações como fraudulentas ou legítimas.
- **Dataset:** Credit Card Fraud Detection
- **ML clássico:** Random Forest, XGBoost, Isolation Forest.
- **Extensão DL:** Autoencoders, MLP com focal loss.
- **Extensão RL:** agente que decide aprovar, bloquear ou verificar.

7. Diagnóstico médico — Breast Cancer Wisconsin

- **Tarefa:** classificar tumores como malignos ou benignos.
- **Dataset:** Breast Cancer Wisconsin
- **ML clássico:** Regressão Logística, SVM, Random Forest.

- **Extensão DL:** MLP, CNNs para imagens médicas.
- **Extensão RL:** tratamento personalizado, triagem ativa.

8. Reconhecimento de atividade humana — UCI HAR

- **Tarefa:** classificar atividades (andar, correr, sentar) a partir de sensores.
- **Dataset:** UCI HAR
- **ML clássico:** Random Forest, SVM.
- **Extensão DL:** CNN, LSTM, CNN-LSTM.
- **Extensão RL:** robótica, reabilitação.

9. Detecção de intrusão em rede — NSL-KDD

- **Tarefa:** classificar tráfego de rede como normal ou ataque.
- **Dataset:** NSL-KDD
- **ML clássico:** Decision Trees, Random Forest, SVM.
- **Extensão DL:** Autoencoders, MLP, redes recorrentes.
- **Extensão RL:** defesa cibernética adaptativa.

10. Classificação de sons ambientais — UrbanSound8K

- **Tarefa:** classificar sons (sirene, latido, música, etc.).
- **Dataset:** UrbanSound8K
- **ML clássico:** MFCC + SVM.
- **Extensão DL:** CNN, CRNN.
- **Extensão RL:** atenção auditiva.

11. Classificação de imagens médicas — Chest X-Ray (Pneumonia)

- **Tarefa:** classificar radiografias de tórax como normais ou com pneumonia.
- **Dataset:** Chest X-Ray Images (Pneumonia)
- **ML clássico:** HOG + SVM, Random Forest.
- **Extensão DL:** CNN, ResNet, DenseNet, transfer learning.
- **Extensão RL:** visão ativa, triagem ativa.

12. Classificação de doenças de pele — Skin Disease Dataset

- **Tarefa:** classificar imagens de lesões cutâneas em diferentes doenças.
- **Dataset:** Skin Disease Dataset
- **ML clássico:** features de textura + SVM.
- **Extensão DL:** CNN, EfficientNet, Vision Transformers.
- **Extensão RL:** triagem dermatológica.

13. Classificação de espécies de folhas — Leaf Classification

- **Tarefa:** classificar 99 espécies de folhas a partir de imagens.
- **Dataset:** Leaf Classification (Kaggle)
- **ML clássico:** extração de forma + SVM.
- **Extensão DL:** CNN, redes siamesas.
- **Extensão RL:** drone agrícola.

14. Classificação de sons de pássaros — BirdSet

- **Tarefa:** classificar espécies de aves a partir de gravações de áudio.
- **Dataset:** BirdSet

- **ML clássico:** MFCC + Random Forest.
 - **Extensão DL:** CNN, CRNN, Transformers para áudio.
 - **Extensão RL:** gravador móvel para detecção de espécies raras.
15. **Classificação de séries temporais — UCR/UEA Archive**
- **Tarefa:** classificar séries temporais em diversos domínios.
 - **Dataset:** UCR/UEA Time Series Classification Archive
 - **ML clássico:** DTW + KNN, Random Forest.
 - **Extensão DL:** ResNet, InceptionTime, TCN.
 - **Extensão RL:** amostragem eficiente de séries temporais.
16. **Classificação de avaliações de produtos — Amazon Reviews**
- **Tarefa:** classificar avaliações em positivas, negativas ou neutras.
 - **Dataset:** Amazon Reviews
 - **ML clássico:** TF-IDF + Regressão Logística.
 - **Extensão DL:** LSTM, BERT.
 - **Extensão RL:** RLHF
17. **Classificação de phishing — Phishing URL Dataset**
- **Tarefa:** classificar URLs como phishing ou legítimas.
 - **Dataset:** Phishing URL Classification
 - **ML clássico:** Random Forest, XGBoost.
 - **Extensão DL:** MLP, redes para dados tabulares.
 - **Extensão RL:** agente de segurança cibernética.
18. **Classificação de qualidade de vinho — Wine Quality**
- **Tarefa:** classificar vinhos como bons ou ruins.
 - **Dataset:** Wine Quality
 - **ML clássico:** Random Forest, SVM.
 - **Extensão DL:** MLP, TabNet.
 - **Extensão RL:** otimização do processo de fermentação.
19. **Classificação de renda — Adult Income (Census)**
- **Tarefa:** prever se a renda anual excede US\$ 50 mil.
 - **Dataset:** Adult Income
 - **ML clássico:** Regressão Logística, Random Forest.
 - **Extensão DL:** MLP, embeddings para variáveis categóricas.
 - **Extensão RL:** políticas públicas.
20. **Classificação de doenças cardíacas — Heart Disease UCI**
- **Tarefa:** classificar pacientes com ou sem doença cardíaca.
 - **Dataset:** Heart Disease UCI
 - **ML clássico:** SVM, Random Forest.
 - **Extensão DL:** MLP, CNN para ecocardiograma.
 - **Extensão RL:** dosagem de medicamentos.
21. **Classificação de diabetes — Pima Indians Diabetes**
- **Tarefa:** prever se uma paciente tem diabetes.
 - **Dataset:** Pima Indians Diabetes
 - **ML clássico:** Regressão Logística, Random Forest.
 - **Extensão DL:** MLP
 - **Extensão RL:** intervenções de estilo de vida.

2 Resumo em Tabela

Projeto	Dataset	ML Clássico	Extensão DL	Extensão RL	
	Roupas	Fashion-MNIST	SVM, RF	CNN	Atenção seletiva
	Imagens	CIFAR-10/100	HOG+SVM	ResNet, ViT	Controle, NAS
Jhonantan	Spam	SMS Spam	Naive Bayes	LSTM, BERT	Bandit contextual
Lygia	Sentimento	IMDB	TF-IDF+SVM	Transformers	RLHF, diálogo
	Notícias	20 Newsgroups	Naive Bayes	BERT	Recomendação
Kaique	Fraude	Credit Card	XGBoost	Autoencoder	Aprovar/bloquear
	Médico	Breast Cancer	SVM, RF	MLP, CNN	Tratamento, triagem
	Atividade	UCI HAR	RF, SVM	CNN-LSTM	Robótica, reabilitação
	Intrusão	NSL-KDD	RF, SVM	Autoencoder	Defesa adaptativa
	Sons	UrbanSound8K	MFCC+SVM	CNN, CRNN	Atenção auditiva
Carlos	Pneumonia	Chest X-Ray	HOG+SVM	CNN, ResNet	Visão ativa
	Pele	Skin Disease	Textura+SVM	EfficientNet	Triagem
Wesley	Folhas	Leaf Classification	Forma+SVM	CNN, Siamesas	Drone agrícola
	Pássaros	BirdSet	MFCC+RF	CNN, CRNN	Gravador móvel
	Séries Temporais	UCR/UEA	DTW+KNN	InceptionTime	Amostragem
Luis G	Amazon	Amazon Reviews	TF-IDF+LR	LSTM, BERT	RLHF
	Phishing	Phishing URL	RF, XGBoost	MLP	Segurança
	Vinho	Wine Quality	RF, SVM	MLP, TabNet	Fermentação
	Renda	Adult Income	LR, RF	MLP	Políticas públicas
	Coração	Heart Disease	SVM, RF	MLP, CNN	Dosagem
	Diabetes	Pima Indians	LR, RF	MLP	Estilo de vida

3 Como Estender de ML para DL e RL

1. **ML → DL:** substitua features manuais por representações aprendidas. Para imagens, CNNs; para texto, Transformers; para séries temporais, RNNs/TCNs.
2. **DL → RL:** reformule o problema como MDP:
 - **Estados:** observações (imagens, sensores, texto).
 - **Ações:** decisões (classificar, bloquear, recomendar, mover).
 - **Recompensas:** acerto, engajamento, lucro, segurança, etc.
3. **Classificação pura → RL:** nem todo problema de classificação vira RL completo. Alternativas:
 - **Bandit contextual:** uma ação por interação, recompensa imediata.
 - **Active learning:** agente escolhe quais amostras rotular.
 - **NAS / hyperparameter tuning:** RL para escolher arquiteturas e hiperparâmetros.
 - **Decisão sequencial:** agente decide onde olhar, qual exame pedir, qual ação tomar.

4 Fontes de Datasets Públicos

- UCI Machine Learning Repository
- Kaggle Datasets
- Hugging Face Datasets
- OpenML
- TensorFlow Datasets
- Papers With Code