

Breast Cancer Classification ICon Project 2024/25

Descrizione del Progetto

Questo progetto ha come obiettivo la creazione di modelli di classificazione per identificare se un tumore al seno è benigno o maligno. Il progetto esplora diversi approcci di apprendimento supervisionato, tra cui alberi decisionali, random forest, Support Vector Machines (SVM), regressione logistica e k-nearest neighbors (KNN), nonché reti neurali e apprendimento bayesiano.

Struttura Progetto

```
root
├── dataset          # Dataset e dataset aumentato
├── img              # Immagini documentazione
│   ├── preprocessing # Analisi preliminare
│   ├── supervised    # Metriche valutazione
│   ├── neuralnetwork # Risultati
│   ├── bayesian      # Grafo e risultati
│   └──
├── scripts          # Contiene tutti i codici sorgente
│   ├── bayesian      # Ragionamento probabilistico
│   ├── dataset_handling # Preprocessing dataset
│   ├── main
│   ├── NNClassifier  # Rete Neurale
│   ├── plotter       # Funzioni per visualizzazione grafici
│   └── supervised_training # Apprendimento supervisionato
```

Dataset

Il [Dataset](#) utilizzato in questo progetto è un dataset pubblico sul cancro al seno.

Obiettivo

L'obiettivo del progetto è allenare, testare e confrontare le prestazioni dei vari modelli di classificazione per prevedere se un tumore al seno è benigno o maligno. L'approccio comprende anche l'uso di tecniche avanzate come l'apprendimento bayesiano e le reti neurali per migliorare l'accuratezza del modello.

Struttura del Progetto

Il progetto è suddiviso nei seguenti passaggi principali:

1. **Pre-elaborazione dei Dati:** Pulizia e preparazione dei dati per l'allenamento.
2. **Allenamento dei Modelli:** Addestramento di vari modelli di machine learning, inclusi Decision Tree, Random Forest, SVM, Logistic Regression, KNN, Neural Networks, e Bayesian Learning.
3. **Valutazione del Modello:** Confronto delle performance dei vari modelli utilizzando metriche di valutazione come l'accuratezza, la precisione, il recall e il F1-score.

4. **Predizioni:** Utilizzo dei modelli per effettuare previsioni su nuovi dati.