

Final Project

Predicting Skin Disease



Supervised Machine Learning Classification



gradio



OpenAI

Oscar Lidheim

Questions



- Can you accurately predict what skin disease a person have?
- How accurate is the prediction only using simple model (12 simple clinical/visual attributes)?
- Could this be a useful application to deploy?

Data

366 patients

6 Skin Diseases (Psoriasis + 5 other)

- 12 Clinical Features (Visual)
- 22 Histopathological Features (Skin sample under microscope)

All data (except age) is on range 0-3

- 0 No presence
- 1-2 low-mid range
- 3 Large presence

Student (Data Engineering Dept)

First program 1998 (Turkey) c. 95% accuracy

More than +10 studies since. 100% accuracy

The screenshot shows a Windows-style dialog box titled "Clinic Features". It contains a list of 12 clinical features, each with five radio buttons corresponding to values: Unknown, 0, 1, 2, and 3. The "Age" field is a text box containing the value "22". The dialog has "Cancel" and "OK" buttons at the bottom.

Feature	Unknown	0	1	2	3
Erythema	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Scaling	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Definite Borders	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Itching	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Koebner Phenomenon	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Polygonal Papules	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Follicular Papules	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Oral Mucosal Involment	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Knee and Elbow Involvement	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scalp Involvement	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Family History	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Age	<input type="text" value="22"/>				

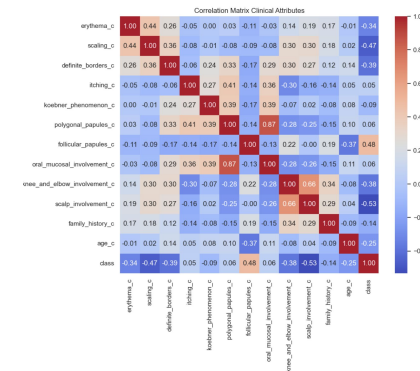
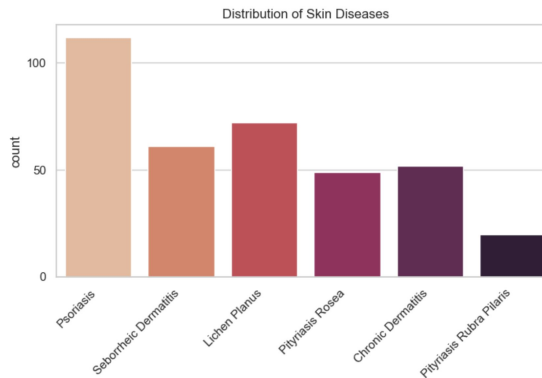
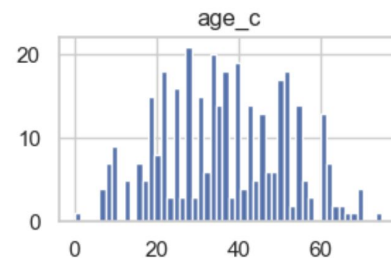
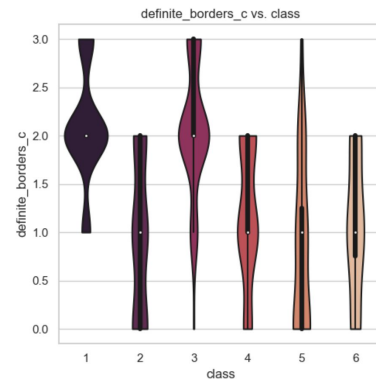
Workflow

- 1) Clean Data
 - a) Split Simple/Full Model
- 2) EDA
- 3) Train 14 models with standard settings (GridsearchCV)
- 4) Tuning two models:

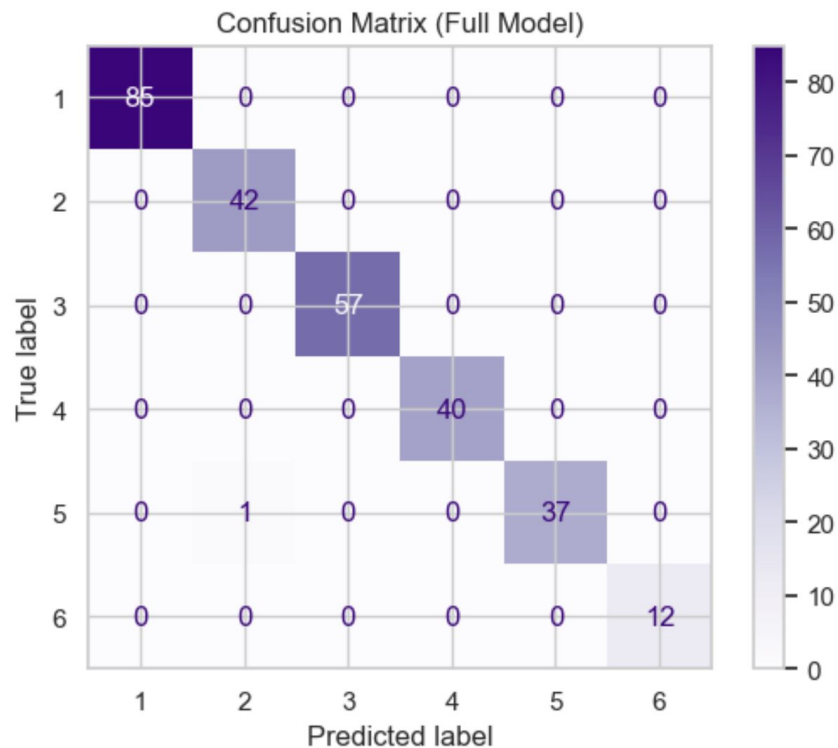
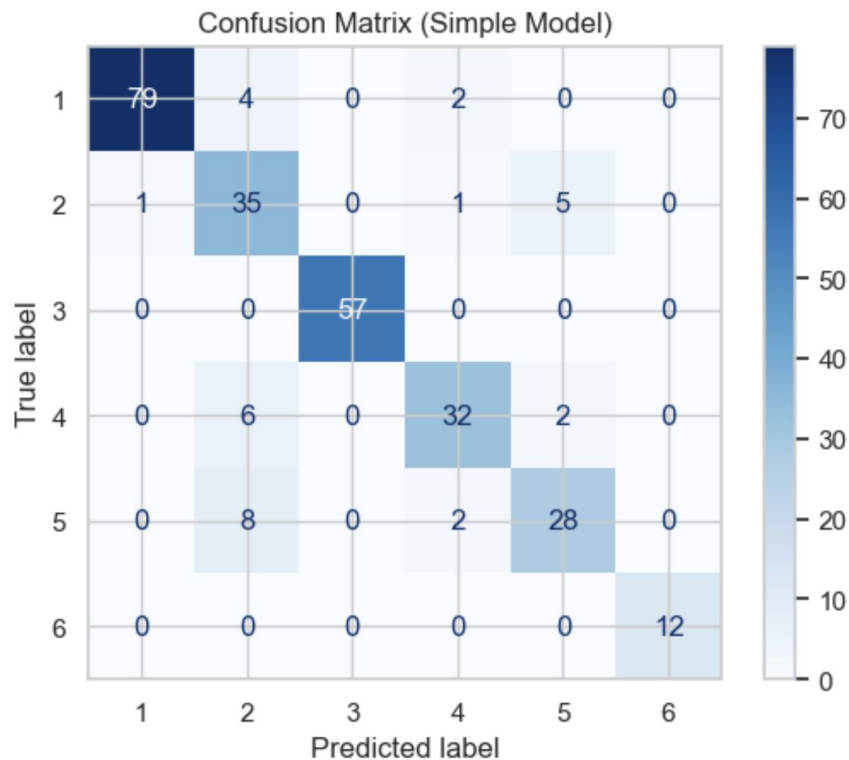
Simple Model: **LinearSVC**

Full Model: **ExtraTreesC**

- 5) Final Evaluation
- 6) Deploy model on web app (Gradio)
 - a) Implement OpenAI API



Confusion Matrix (Training)



Final Evaluation

Simple model (LinearSVC):

Accuracy 88%

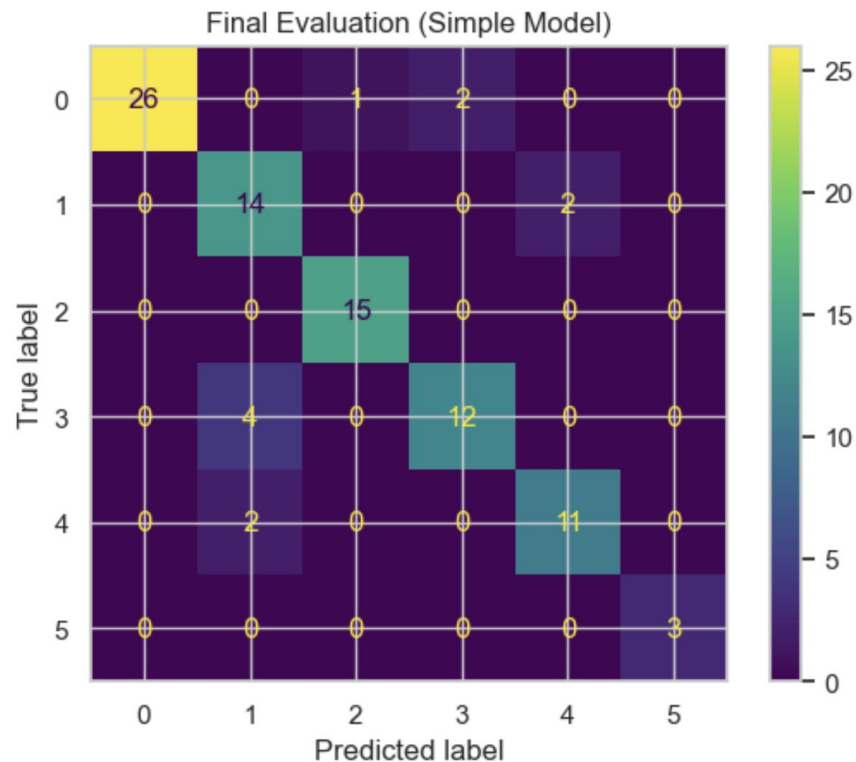
Precision 89%

Recall 88%

F1 88%

Full model (ExtraTrees):

100%



Challenges & Learnings

Challenges

- Visual Studio Crash (4hrs)
- Gradio took most time (minor changes)
- Keeping track of saved models when running program

Learnings

- Saving, prediction and deployment of ML models
- Using new library Gradio
 - Working with visuals/UI



Thank You