COMP 472 Assignment 1

Team Waluigi
Daniel Wiktorczyk 40060894
Waqar Qureshi 40055526

Team Waluigi
Contributions

Waqar

- Own half of the Models
- Start the project by Implementing Skilearn for each model

Daniel

- Own half of the Models
- Part 1 Code and Analysis
- Implement GridSearchCV

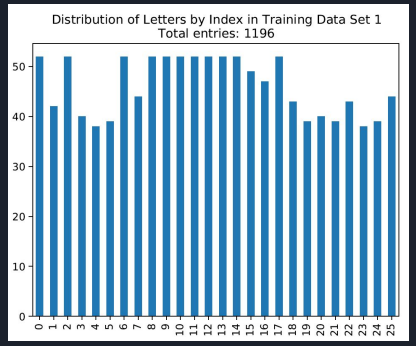
Both Members:

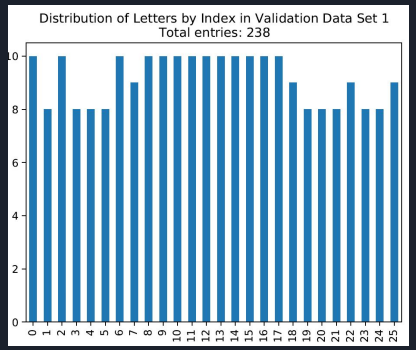
- ✔ Review each other's work during live session
- ✔ Prepare for presentation
- ✔ Understand the entire code
- ✔ Understand all these metrics
- ✔ **Equal Participation**

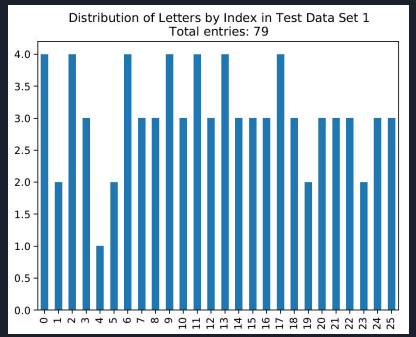
Part 1: Dataset Analysis

Data Set 1: Latin Alphabet

- 26 classes [a-z] (Higher)
- Training set to Validation set composition
  - 1196:238 , or 83/17 split, ~80/20 convention
  - More or less distributed equally
- Test set 79 entries
  - 7% of training data, low sample set, => less accurate measures
  - Some entry outliers, e.g. letter ‘E’ only represented once

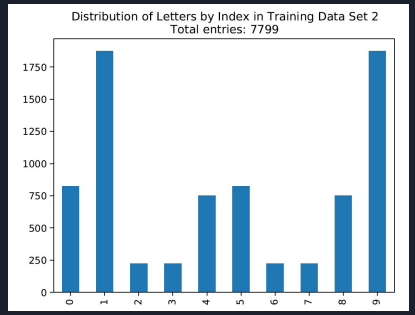


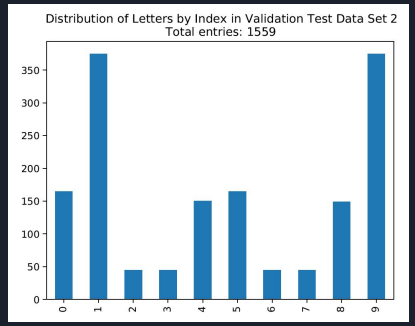


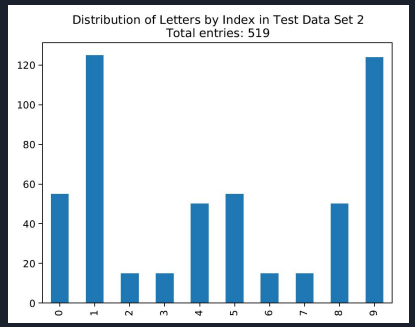


Data Set 2: Greek Alphabet

- 26 classes [pi-xi] (Lower)
- Training set to Validation set composition
  - 7799:1559 , or 80/20 split, perfect convention!
  - NOT distributed equally!
    - Alpha(1) represented over 1750 times, but Beta(2) around 250
  - Composition remains equal! Good
- Test set 512 entries
  - 7% of training data, low sample set, => less accurate measure
  - Same composition as training and validation sets! Good, representative
    - => more accurate measure







Part 2: Models

GNB-DS1

Confusion Matrix

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	3	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
C	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
G	0	1	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
K	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0
L	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1
M	0	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
N	1	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0
P	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0
S	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0
Z	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2

Precision

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0.5	0.14286	1	0.5	1	1	0.75	0	1	1	0.5	1	0.4	1	1	1	1	0.42857	1	1	1	0.66667	0.75	0.66667	1	0.66667

Recall

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0.75	0.5	0.75	0.66667	1	0.5	0.75	0	1	0.5	0.33333	0.75	0.66667	0.25	1	0.66667	1	0.75	0.66667	1	1	0.66667	1	1	0.66667	0.66667

F1-Measures

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0.6	0.22222	0.85714	0.57143	1	0.66667	0.75	0	1	0.66667	0.4	0.85714	0.5	0.4	1	0.8	1	0.54545	0.8	1	1	0.66667	0.85714	0.8	0.8	0.66667

Accuracy	0.6962
F1-Macro	0.70874
F1-Weighted	0.69961

GNB-DS2

Confusion Matrix										
	pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
pi	31	5	1	9	7	1	0	0	1	0
alpha	5	112	1	6	0	1	0	0	0	0
beta	0	0	6	0	0	0	0	0	4	5
sigma	1	0	0	14	0	0	0	0	0	0
gamma	3	0	0	6	18	2	4	0	1	16
delta	0	0	0	1	0	33	7	0	0	14
lambda	3	1	0	0	0	0	11	0	0	0
omega	0	1	0	0	0	0	0	14	0	0
mu	2	2	7	0	0	0	5	0	29	5
xi	4	1	1	0	4	7	15	0	0	92

Precision									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.632653061	0.918033	0.375	0.388889	0.62069	0.75	0.261905	1	0.828571	0.69697

Recall									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.563636364	0.896	0.4	0.933333	0.36	0.6	0.733333	0.933333	0.58	0.741935

F1-Measures									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.596153846	0.906883	0.387097	0.54902	0.455696	0.666667	0.385965	0.965517	0.682353	0.71875

Accuracy	0.693642
F1-Macro	0.63141
F1-Weighted	0.699724

Summary of Base GNB (Default)

Performance is up to par with the better models studied

- Performance can be attributed to the consideration of each feature in calculations
- Flaws can be attributed to the fact that features are equally considered, the corner pixels may be considered just as important as the core ones.
- Performances are similar across both alphabets because of the weighting issue despite lower number of classes.
- The model seems to be unsuited to this dataset due to our knowledge of many pixels being useless to the determination of greek and latin letter.

Base-DT-DS1

Confusion Matrix

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
C	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
D	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
E	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
I	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
J	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
K	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	1	0	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
S	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0
W	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
Y	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
Z	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0

Precision

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	0	0.6	0.5	0.5	1	1	0.33333	1	0.66667	0	1	0.33333	0.375	0.5	0.4	1	0.66667	0.5	0.25	1	0.5	0.5	0.5	0.5	0

Recall

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	0	0.75	0.33333	1	0.5	0.5	0.33333	0.66667	0.5	0	0.75	0.33333	0.75	1	0.66667	0.33333	0.5	0.33333	0.5	1	0.66667	0.33333	1	0.66667	0

F1-Measures

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	0	0.66667	0.4	0.66667	0.66667	0.66667	0.33333	0.8	0.57143	0	0.85714	0.33333	0.5	0.66667	0.5	0.5	0.57143	0.4	0.33333	1	0.57143	0.4	0.66667	0.57143	0

Accuracy	0.55696
F1-Macro	0.52473
F1-Weighted	0.54129

Base-DT-DS2

Confusion Matrix										
	pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
pi	42	0	4	1	2	1	2	0	0	3
alpha	3	109	0	2	2	1	1	1	5	1
beta	0	0	8	0	1	2	0	0	3	1
sigma	0	4	0	9	1	0	0	1	0	0
gamma	4	2	1	0	27	7	0	1	1	7
delta	1	0	1	1	6	42	0	0	0	4
lambda	2	0	0	0	2	0	7	0	2	2
omega	0	1	0	0	0	0	0	13	1	0
mu	0	0	2	0	0	1	1	0	45	1
xi	2	1	3	0	12	8	0	0	4	94

Precision									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.777777778	0.931623932	0.421053	0.692308	0.509434	0.677419	0.636364	0.8125	0.737705	0.831858
Recall									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.763636364	0.872	0.533333	0.6	0.54	0.763636	0.466667	0.866667	0.9	0.758065
F1-Measures									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.770642202	0.900826446	0.470588	0.642857	0.524272	0.717949	0.538462	0.83871	0.810811	0.793249

Accuracy	0.76300578
F1-Macro	0.700836556
F1-Weighted	0.764840187

Summary of Base Decision Tree

Performance is among the lowest observed

- This can be attributed to how decision trees are built and how influential decisions at the top are.
- With 32x32 = 1024 features to consider, a decision tree may often arrive with a bias to a certain decision before reaching a leaf node having considered all features.
- Entropy is used in the base DT, which prioritizes maximum information gain, versus the gini impurity which calculates the divergence at particular nodes.
- Better performance in set 2, since there are less classification groups. We can observe that we are getting better results simply because we are classifying popular groups correct often and less popular groups incorrectly. A symptom of the top heavy nature of this model
- This model is not well suited for the data to be analyzed.

Best-DT-DS1

Confusion Matrix	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	3	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
B	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
C	0	0	2	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
E	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
G	0	0	1	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
I	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
K	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0
L	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0
O	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
Z	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1

Precision																										
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
0.75	1	0.66667	1	1	1	0.4	1	0.75	0.66667	0	1	0.5	0.5	0.4	0.5	0.75	0.5	1	0.16667	1	0	0.5	0.33333	0.5	0.33333	

Recall																										
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
0.75	0.5	0.5	0.33333	1	0.5	0.5	0.66667	1	0.5	0	1	0.66667	0.5	0.66667	0.66667	1	0.5	0.66667	0.5	0.33333	0	0.33333	1	0.33333	0.33333	

F1-Measures																										
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
0.75	0.66667	0.57143	0.5	1	0.66667	0.44444	0.8	0.85714	0.57143	0	1	0.57143	0.5	0.5	0.57143	0.85714	0.5	0.8	0.25	0.5	0	0.4	0.5	0.4	0.33333	

Accuracy	0.55696
F1-Macro	0.55812
F1-Weighted	0.55427

Best Parameters : {'class_weight': None, 'criterion': 'gini', 'max_depth': None, 'min_impurity_decrease': 0, 'min_samples_split': 2}

Best-DT-DS2

Confusion Matrix										
	pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
pi	45	1	0	0	4	2	0	0	0	3
alpha	1	110	1	5	1	1	1	0	4	1
beta	3	2	6	0	0	1	0	0	1	2
sigma	0	1	0	13	0	1	0	0	0	0
gamma	4	0	0	1	26	4	0	1	3	11
delta	2	0	3	0	5	38	0	0	0	7
lambda	1	0	0	0	2	0	7	0	1	4
omega	0	0	0	0	0	0	0	15	0	0
mu	1	1	0	0	2	1	0	0	45	0
xi	3	1	1	0	9	6	3	0	2	99

Precision									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.75	0.948276	0.545455	0.684211	0.530612	0.703704	0.636364	0.9375	0.803571	0.779528

Recall									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.818181818	0.88	0.4	0.866667	0.52	0.690909	0.466667	1	0.9	0.798387

F1-Measures									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.782608696	0.912863	0.461538	0.764706	0.525253	0.697248	0.538462	0.967742	0.849057	0.788845

Accuracy	0.77842
F1-Macro	0.728832
F1-Weighted	0.77653

Best Parameters : {'class_weight': 'balanced', 'criterion': 'entropy', 'max_depth': None, 'min_impurity_decrease': 0, 'min_samples_split': 3}

Summary of **Best** Decision Tree

Performance at par with the Base Decision Tree

- Despite calculating the ideal classifying parameters, it is observed that the performance of this model is equivalent to the base model.
- This is a direct consequence of this model being unsuited to this problem.

Chosen parameters

- No maximum depth: Important due to the number of features
- Lowest samples required to split out of our options, leads to earlier splitting and makes the tree more flexible.
- Impurity decrease is chosen to be zero, which leads to any impurity decrease being an acceptable split
- Class weight chosen to be zero as the dataset is balanced, so there is no need for a lean.

PER-DS1

Confusion Matrix

Act\Pre	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
B	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	1	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
K	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	1	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
T	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	0
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
Z	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0

Precision

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0.75	0.5	1	1	1	0.66667	0.75	0.33333	1	1	0.33333	1	0.66667	1	1	1	1	0.75	1	0.5	1	0.5	0.5	0.66667	1	1

Recall

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0.75	1	1	1	1	1	0.75	0.33333	1	0.75	0.33333	1	0.66667	0.75	1	0.66667	1	0.75	0.66667	0.5	1	0.33333	1	1	1	0.33333

F1-Measures

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0.75	0.66667	1	1	1	0.8	0.75	0.33333	1	0.85714	0.33333	1	0.66667	0.85714	1	0.8	1	0.75	0.8	0.5	1	0.4	0.66667	0.8	1	0.5

Accuracy	0.78481
F1-Macro	0.77811
F1-Weight	0.78342

PER-DS2

Confusion Matrix

Act\Pre	pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
pi	53	0	0	0	0	0	0	0	0	2
alpha	2	120	0	2	0	0	0	0	1	0
beta	3	0	8	0	1	0	0	0	1	2
sigma	0	0	0	14	1	0	0	0	0	0
gamma	8	1	0	3	27	0	0	0	1	10
delta	0	0	2	1	2	43	0	0	0	7
lambda	1	4	1	0	1	0	4	0	1	3
omega	0	0	0	0	0	0	0	15	0	0
mu	9	1	0	0	0	0	0	0	37	3
xi	5	0	1	0	6	2	0	0	3	107

Precision

pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.65432	0.95238	0.66667	0.7	0.71053	0.95556	1	1	0.84091	0.79851

Recall

pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.96364	0.96	0.53333	0.93333	0.54	0.78182	0.26667	1	0.74	0.8629

F1-Measures

pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.77941	0.95618	0.59259	0.8	0.61364	0.86	0.42105	1	0.78723	0.82946

Accuracy	0.82466
F1-Macro	0.76396
F1-Weight	0.81848

Summary of Perceptron

- Best base default performance (78%, 82% accuracy)
- Very well done for a single perceptron!
- Considers all features equally initially, adjusts weights accordingly
    - Many features are unimportant in letters (e.g. top left pixel), so we can expect this to be a benefit

Set 2 performance was better

    - Possibly due to the less amount of output classes (10 vs 26)
    - F1 macro lower because of some classes less represented, less performant
- We can expect the MLP to perform better…?

Base-MLP-DS1

Confusion Matrix

Act\Pre	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
C	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
E	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
K	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
L	0	0	1	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	1	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0
V	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0
W	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	0	0	0	0	0	0	0
Z	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Precision

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0.75	0	0.571429	0	0	0	0.75	0.5	0.333333	0.4	0.4	0.4	0.5	0.375	0.4	0.5	1	0.230769	0	0	0	0	0.666667	0	0	0

Recall

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0.75	0	1	0	0	0	0.75	0.333333	1	0.5	0.666667	0.5	0.666667	0.75	0.666667	0.333333	1	0.75	0	0	0	0	0.666667	0	0	0

F1-Measures

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0.75	0	0.727273	0	0	0	0.75	0.4	0.5	0.444444	0.5	0.444444	0.571429	0.5	0.5	0.4	1	0.352941	0	0	0	0	0.666667	0	0	0

Accuracy	0.455696
F1-Macro	0.3272
F1-Weigh	0.3733

Base-MLP-DS2

Confusion Matrix										
Act\Pre	pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
pi	47	1	1	0	0	2	0	0	2	2
alpha	3	119	0	1	0	1	0	0	1	0
beta	0	0	8	0	0	1	0	0	5	1
sigma	2	0	0	12	0	1	0	0	0	0
gamma	6	1	0	0	29	2	0	0	3	9
delta	0	0	0	0	1	50	0	0	0	4
lambda	1	3	1	0	1	0	6	0	2	1
omega	0	0	0	0	0	0	0	15	0	0
mu	0	1	0	0	0	0	0	0	45	4
xi	2	0	0	0	7	5	0	0	1	109

Precision									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.77049	0.952	0.8	0.92308	0.76316	0.80645	1	1	0.76271	0.83846
Recall									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.85455	0.952	0.53333	0.8	0.58	0.90909	0.4	1	0.9	0.87903

F1-Measures									
pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.81034	0.952	0.64	0.85714	0.65909	0.8547	0.57143	1	0.82569	0.85827

Accuracy	0.84778
F1-Macro	0.80287
F1-Weight	0.84252

Summary of Base MLP

- Set1 performed very poorly in comparison with all other models! (46% acc)
  - Many output classifiers to blame, along with lower sample set
  - Default 100 hidden nodes with only 200 iterations, back propagation error likely still exists
  - We tested with more iterations for fun (2000), we got 80% accuracy!
- Set2 performance is better (85% acc)
  - likely thanks to less output classes (10 instead of 26 outputs) and higher sample set
  - F1 macro lower because of some classes less represented, less performant

Best-MLP-DS1

Confusion Matrix

Act\Pre	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
K	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	1	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	3	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
U	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
Z	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2

Precision

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	1	1	1	1	1	0.75	0.6	1	1	0.5	1	1	1	1	0.75	0.75	1	1	0.66667	1	1	1	1	1	1

Recall

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	1	1	1	1	0.5	0.75	1	1	0.75	0.66667	1	0.66667	1	1	1	1	0.75	0.66667	1	1	1	1	1	1	0.66667

F1-Measures

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	1	1	1	1	0.66667	0.75	0.75	1	0.85714	0.57143	1	0.8	1	1	0.85714	0.85714	0.85714	0.8	0.8	1	1	1	1	1	0.8

Accuracy	0.89873
F1-Macro	0.89872
F1-Weight	0.89997

Best Parameters : {'activation': 'identity', 'hidden_layer_sizes': 50, 'solver': 'adam'}

Best-MLP-DS2

Confusion Matrix

Act\Pre	pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
pi	50	0	0	1	1	2	0	0	0	1
alpha	2	122	0	0	0	0	0	0	1	0
beta	0	0	9	0	1	0	0	0	3	2
sigma	1	1	0	13	0	0	0	0	0	0
gamma	4	1	0	0	43	1	0	0	0	1
delta	0	0	1	0	1	51	0	0	0	2
lambda	0	2	1	0	1	0	11	0	0	0
omega	0	0	0	0	0	0	0	15	0	0
mu	0	1	0	0	0	0	0	0	47	2
xi	1	0	1	0	7	5	0	0	4	106

Precision

pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.86207	0.96063	0.75	0.92857	0.7963	0.86441	1	1	0.85455	0.92982

Recall

pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.90909	0.976	0.6	0.86667	0.86	0.92727	0.73333	1	0.94	0.85484

F1-Measures

pi	alpha	beta	sigma	gamma	delta	lambda	omega	mu	xi
0.88496	0.96825	0.66667	0.89655	0.82692	0.89474	0.84615	1	0.89524	0.89076

Accuracy	0.89981
F1-Macro	0.87702
F1-Weight	0.89907

Param grid tested:

{'activation': ['logistic', 'tanh', 'relu', 'identity'], 'hidden_layer_sizes': [50, (25, 25)], 'solver': ['adam', 'sgd']}

Chosen best params:

{'activation': 'tanh', 'hidden_layer_sizes': 50, 'solver': 'adam'}

Best Parameters : {'activation': 'tanh', 'hidden_layer_sizes': 50, 'solver': 'adam'}

Summary of **Best** MLP

- Set1 and 2
  - 90% Accuracy,
  - Set2 Lower F2 Macro because of class representation
- Chosen parameters

```
param_grid = {
    'activation': ['logistic', 'tanh', 'relu', 'identity'],
    'hidden_layer_sizes': [(50), (25, 25)],
    'solver': ['adam', 'sgd']
}
GridSearchCV chose…
Set 1 "{'activation': 'identity', 'hidden_layer_sizes': 50, 'solver': 'adam'}"
Set 2 "{'activation': 'tanh',     'hidden_layer_sizes': 50, 'solver': 'adam'}"
```

- On repeat runs, GridSearchCV jumps between [identity, tanh] => equivalent performance
  - Random initialization to blame
- Simpler architecture favoured, 50 vs (25, 25) => less error given small iteration amount

Thank You!

- Questions?
- New Dataset!
- **W**aaaaaaaah!
- https://github.com/TeamWaluigi
  - Includes all Input and Output files
  - README, Repo and Google slides link