**EssentiaLL**

1. Best 10 features

|  |  |  |
| --- | --- | --- |
| English | Chinese | Turkish |
| barkbands\_spread.mean  spectral\_complexity.mean  spectral\_centroid.mean  spectral\_energyband\_high.mean  spectral\_complexity.stdev  spectral\_rolloff.mean  melbands\_spread.mean  spectral\_skewness.stdev  zerocrossingrate.mean  Dissonance.stdev | silence\_rate\_20dB.stdev  silence\_rate\_60dB.stdev  hfc.stdev  spectral\_energyband\_high.stdev  silence\_rate\_60dB.mean  silence\_rate\_30dB.stdev  spectral\_energyband\_high.mean  hfc.mean  pitch\_salience.stdev  spectral\_energyband\_low.mean | silence\_rate\_30dB.stdev  spectral\_centroid.mean  spectral\_skewness.stdev  melbands\_spread.mean  spectral\_kurtosis.stdev  silence\_rate\_60dB.stdev  barkbands\_spread.mean  zerocrossingrate.mean  barkbands\_flatness\_db.mean  spectral\_kurtosis.mean |

1. Best 3 algorithms

English => MLP, RF, KNN

KNN: acc = 0.66, +-0.02, values: [0.63 0.67 0.68 0.65 0.68 0.68 0.62 0.68 0.63 0.68]

Linear SVM: acc = 0.62, +-0.03, values: [0.64 0.61 0.62 0.57 0.64 0.6 0.62 0.6 0.67 0.62]

RBF SVM: acc = 0.60, +-0.03, values: [0.58 0.63 0.62 0.57 0.57 0.62 0.56 0.63 0.61 0.6 ]

Gaussian Process: acc = 0.44, +-0.05, values: [0.43 0.44 0.45 0.58 0.41 0.48 0.41 0.38 0.41 0.44]

Neural Net: acc = 0.68, +-0.02, values: [0.68 0.67 0.68 0.67 0.7 0.69 0.67 0.66 0.68 0.72]

Naive Bayes: acc = 0.48, +-0.02, values: [0.49 0.51 0.46 0.45 0.45 0.48 0.5 0.48 0.46 0.5 ]

Random Forest: acc = 0.68, +-0.03, values: [0.7 0.71 0.68 0.62 0.68 0.68 0.71 0.69 0.65 0.68]

Chinese => KNN, RBF, RF

KNN: acc = 0.49, +-0.03, values: [0.43 0.57 0.48 0.49 0.51 0.47 0.51 0.46 0.47 0.49]

Linear SVM: acc = 0.33, +-0.06, values: [0.25 0.28 0.3 0.42 0.33 0.23 0.34 0.43 0.33 0.35]

RBF SVM: acc = 0.48, +-0.05, values: [0.42 0.47 0.45 0.54 0.54 0.37 0.53 0.46 0.44 0.53]

Gaussian Process: acc = 0.22, +-0.03, values: [0.22 0.2 0.25 0.28 0.21 0.17 0.23 0.23 0.2 0.24]

Neural Net: acc = 0.46, +-0.07, values: [0.3 0.48 0.44 0.59 0.45 0.43 0.53 0.48 0.39 0.48]

Naive Bayes: acc = 0.27, +-0.04, values: [0.25 0.23 0.25 0.33 0.22 0.28 0.27 0.35 0.29 0.26]

Random Forest: acc = 0.47, +-0.06, values: [0.34 0.52 0.46 0.52 0.51 0.45 0.5 0.52 0.42 0.51]

Turkish => GP, MLP, RF

KNN: acc = 0.80, +-0.05, values: [0.82 0.82 0.75 0.78 0.72 0.78 0.75 0.87 0.85 0.83]

Linear SVM: acc = 0.81, +-0.03, values: [0.85 0.79 0.77 0.81 0.75 0.8 0.85 0.82 0.82 0.84]

RBF SVM: acc = 0.71, +-0.06, values: [0.8 0.73 0.62 0.61 0.73 0.76 0.68 0.79 0.64 0.74]

Gaussian Process: acc = 0.89, +-0.02, values: [0.9 0.86 0.87 0.87 0.88 0.89 0.87 0.94 0.89 0.92]

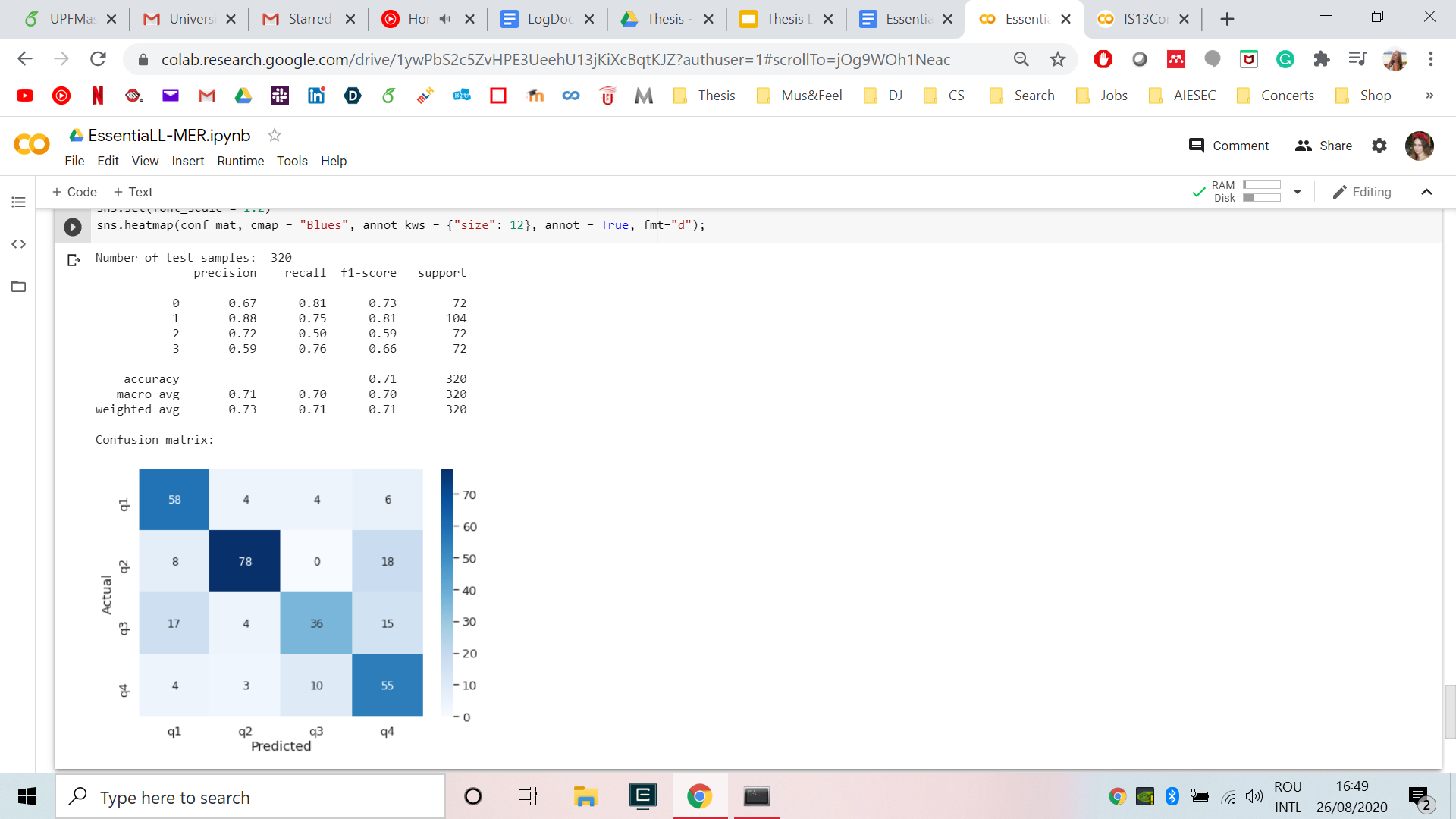
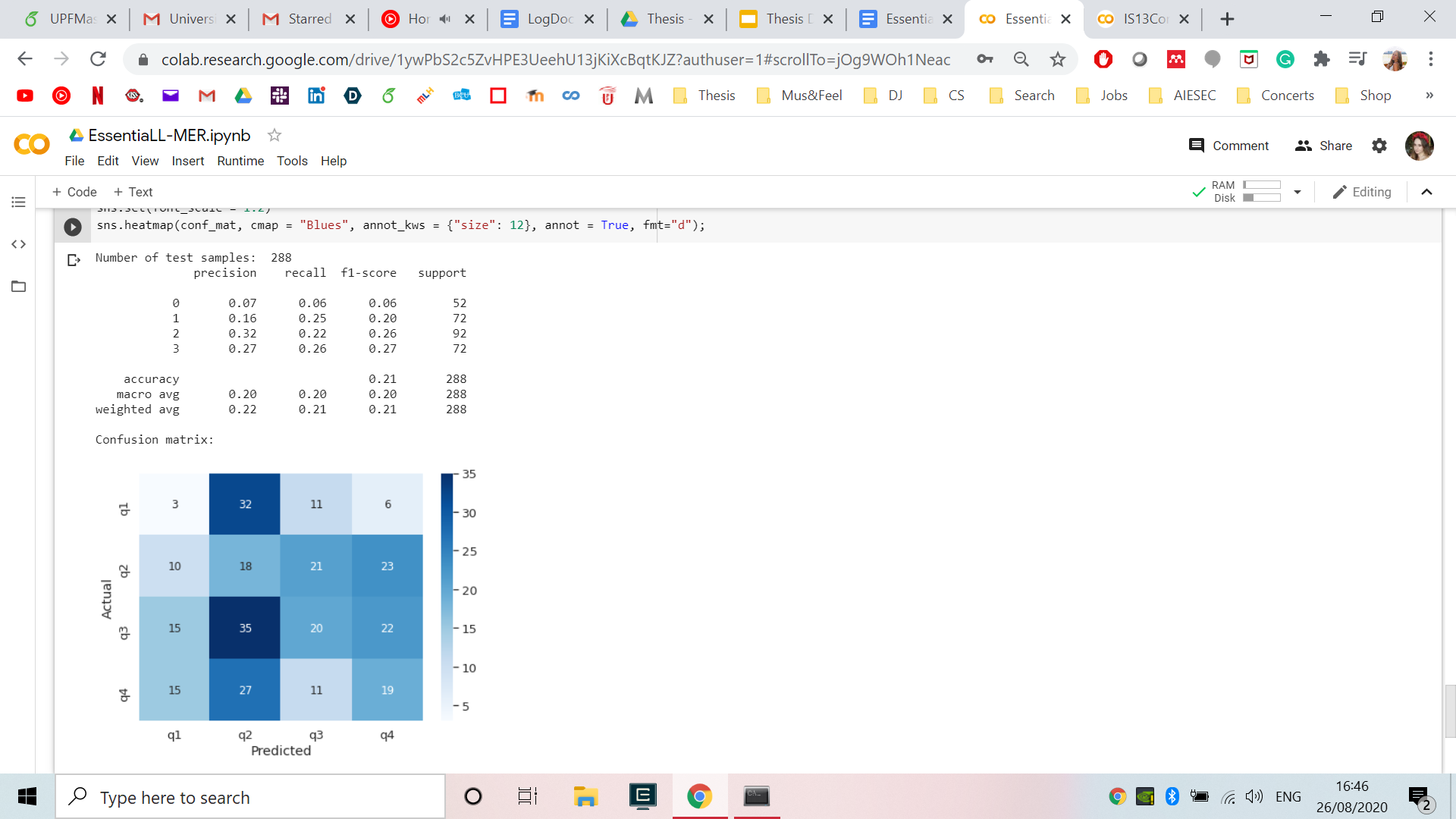
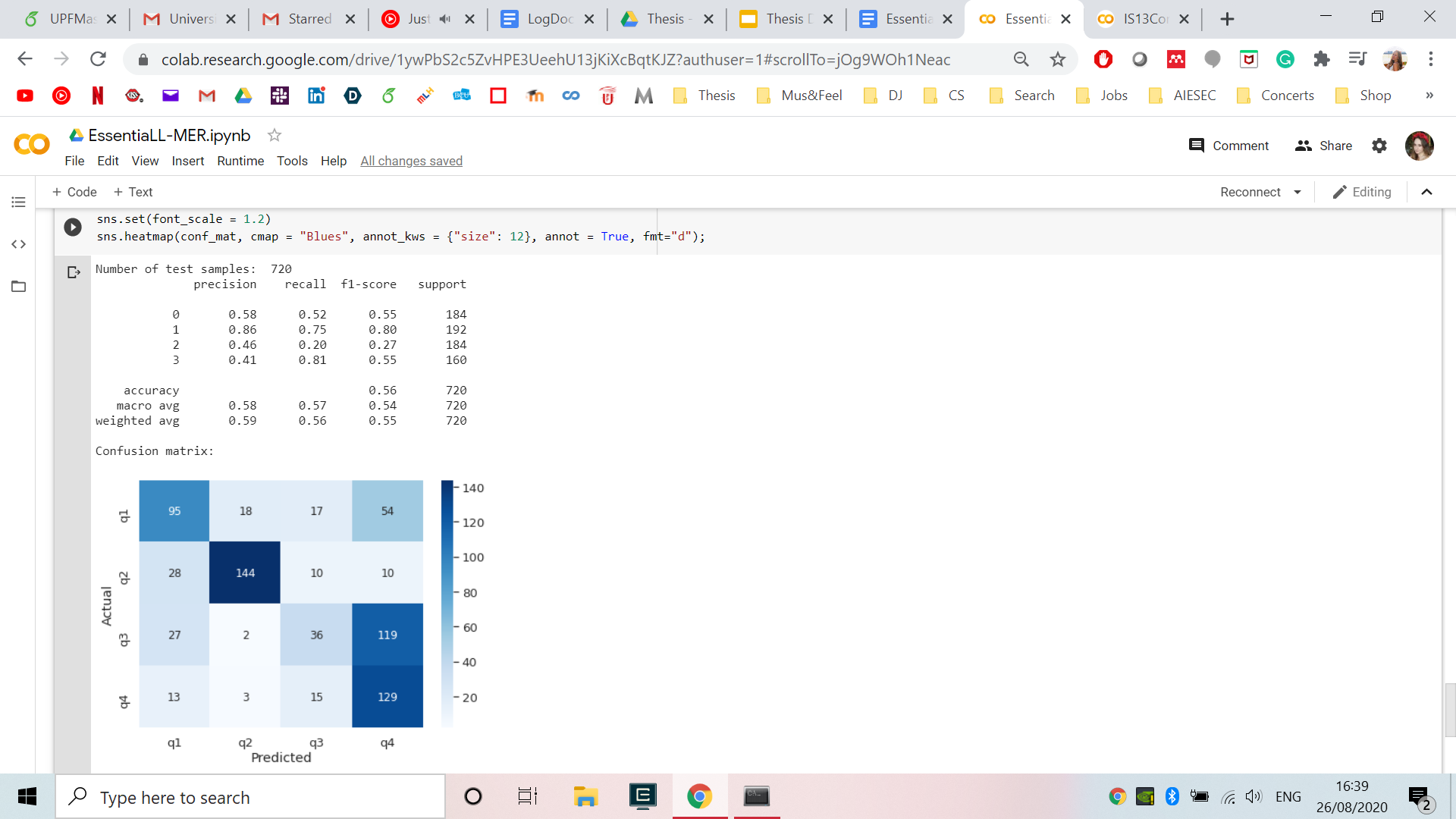
Neural Net: acc = 0.89, +-0.04, values: [0.9 0.86 0.82 0.87 0.87 0.87 0.92 0.94 0.88 0.93]

Naive Bayes: acc = 0.68, +-0.03, values: [0.7 0.63 0.66 0.68 0.66 0.63 0.68 0.7 0.68 0.74]

Random Forest: acc = 0.84, +-0.03, values: [0.86 0.82 0.8 0.84 0.83 0.83 0.79 0.86 0.87 0.89]

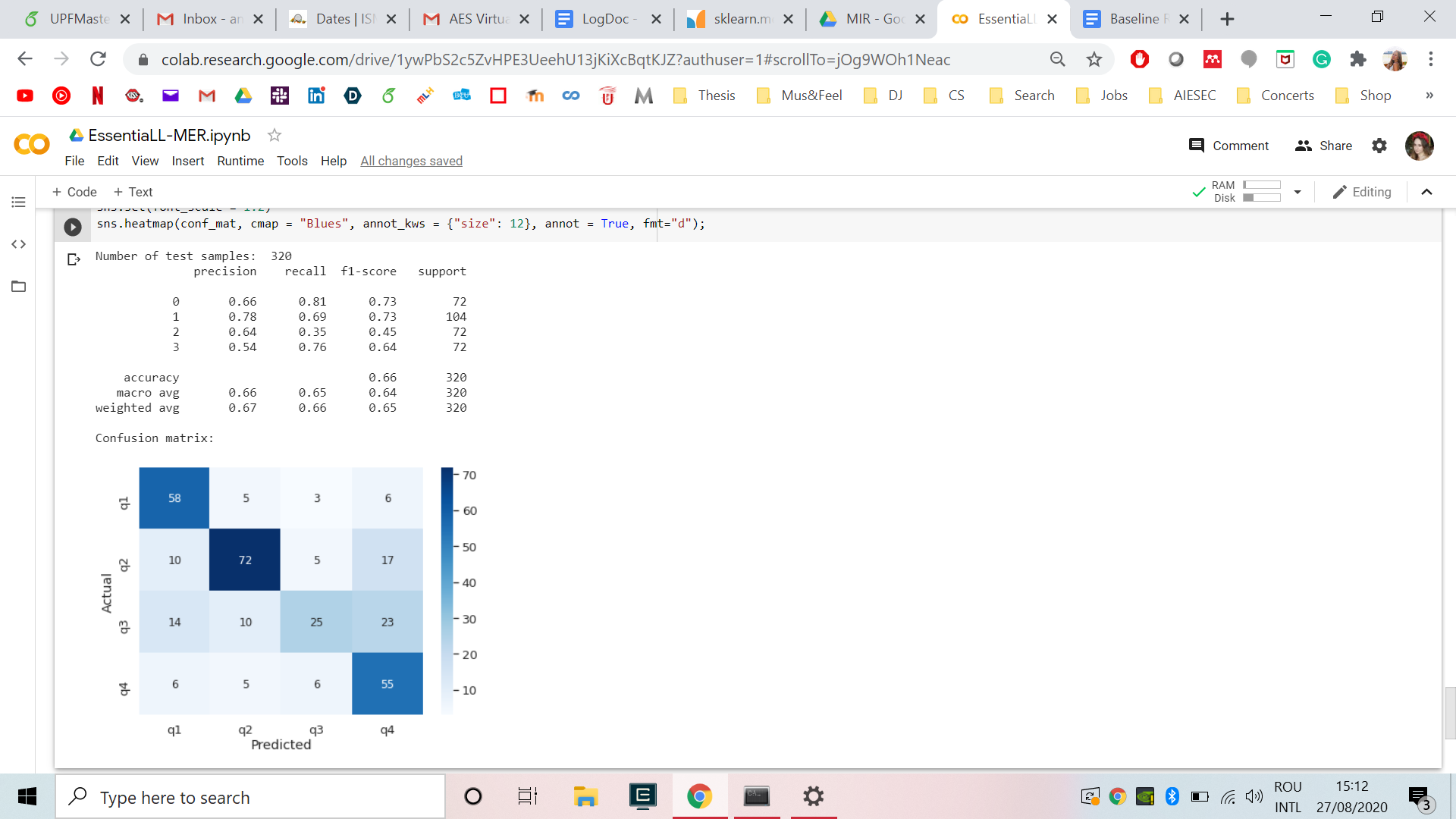
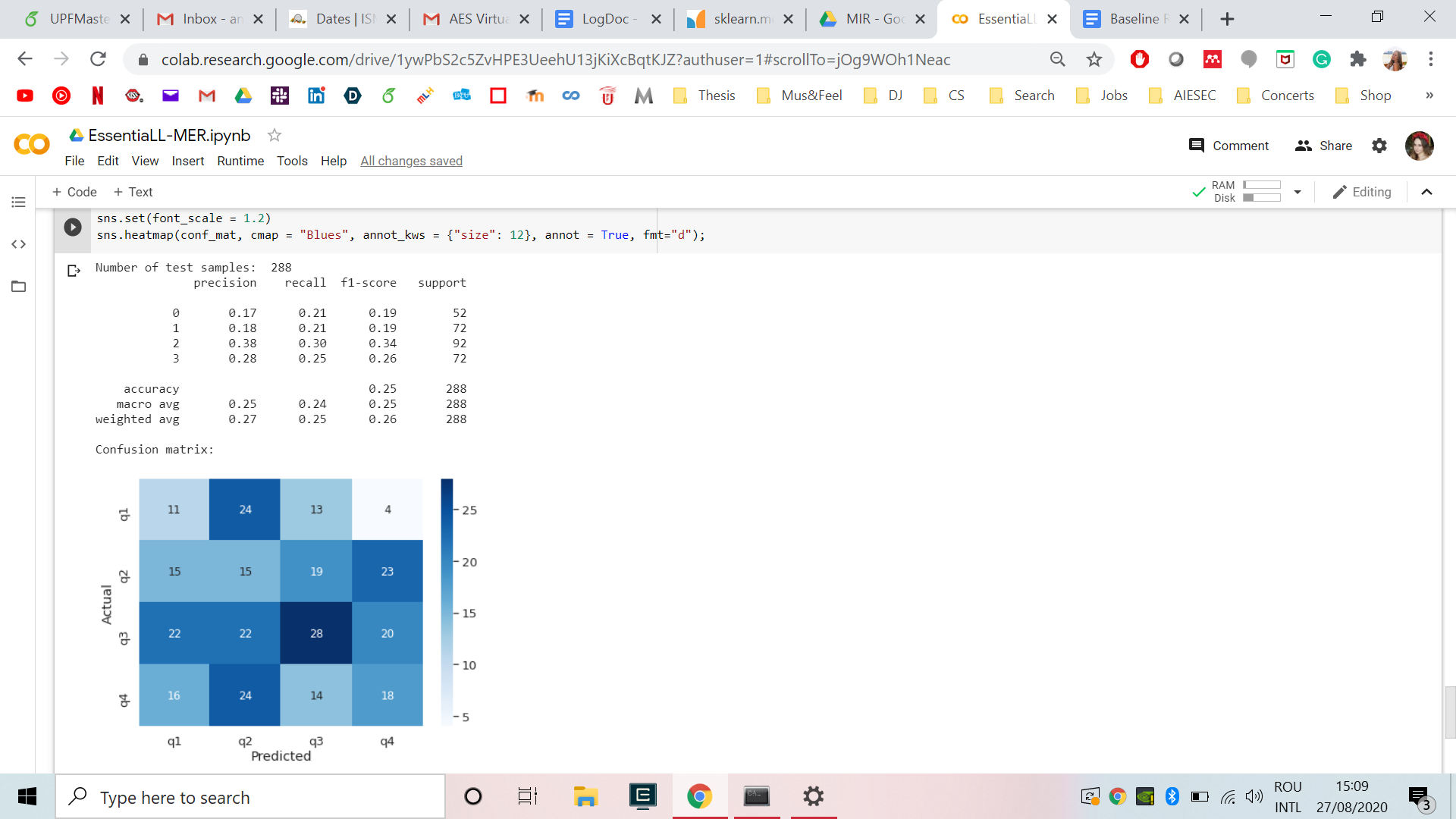
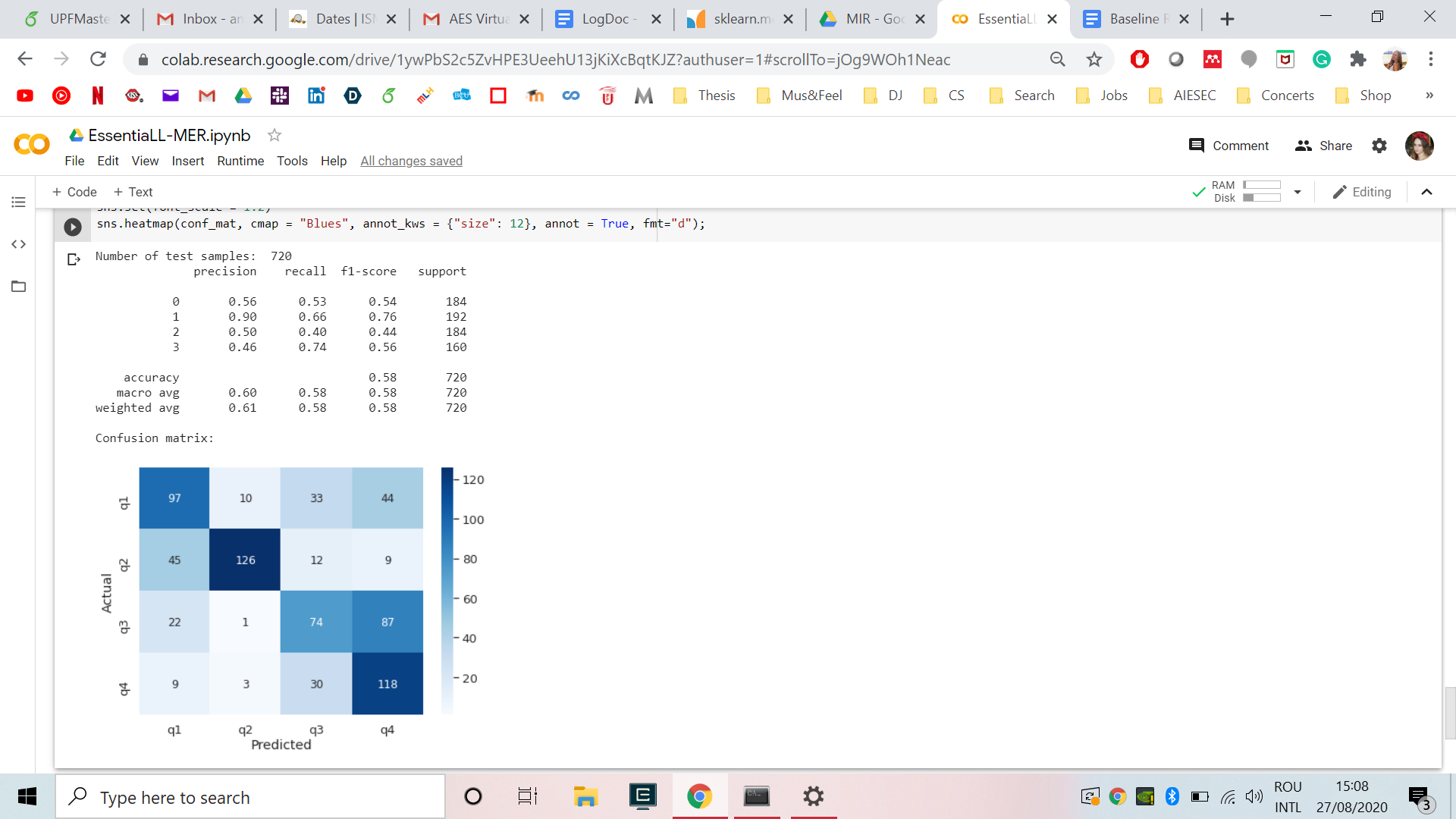
1. MLP results

|  |  |  |  |
| --- | --- | --- | --- |
|  | English | Chinese | Turkish |
| Precision | 59% | 22% | 73% |
| Recall | 56% | 21% | 71% |
| F-score | 55% | 21% | 71% |
| Accuracy | 56% | 21% | 71% |

****

1. RF results

|  |  |  |  |
| --- | --- | --- | --- |
|  | English | Chinese | Turkish |
| Precision | 61% | 27% | 67% |
| Recall | 58% | 25% | 66% |
| F-score | 58% | 26% | 65% |
| Accuracy | 58% | 25% | 66% |

****

**IS13ComParE**

1. Best 10 features

|  |  |  |
| --- | --- | --- |
| English | Chinese | Turkish |
| logHNR\_sma\_stddev  pcm\_fftMag\_spectralRollOff90.0\_sma\_amean  pcm\_fftMag\_spectralVariance\_sma\_amean  mfcc\_sma[1]\_amean  pcm\_fftMag\_spectralRollOff75.0\_sma\_amean  pcm\_fftMag\_spectralFlux\_sma\_amean  audspec\_lengthL1norm\_sma\_amean  pcm\_fftMag\_spectralCentroid\_sma\_amean  logHNR\_sma\_amean  pcm\_zcr\_sma\_amean | logHNR\_sma\_de\_stddev  logHNR\_sma\_stddev  pcm\_fftMag\_fband1000-4000\_sma\_stddev  F0final\_sma\_stddev  pcm\_fftMag\_spectralHarmonicity\_sma\_de\_stddev  pcm\_fftMag\_spectralSkewness\_sma\_de\_stddev  pcm\_fftMag\_fband1000-4000\_sma\_de\_stddev  pcm\_fftMag\_fband1000-4000\_sma\_amean  mfcc\_sma\_de[1]\_stddev  shimmerLocal\_sma\_amean | logHNR\_sma\_stddev  pcm\_fftMag\_spectralCentroid\_sma\_amean  pcm\_fftMag\_spectralRollOff75.0\_sma\_amean  pcm\_fftMag\_spectralRollOff90.0\_sma\_amean  pcm\_fftMag\_spectralRollOff50.0\_sma\_de\_stddev  pcm\_fftMag\_spectralVariance\_sma\_amean  pcm\_zcr\_sma\_amean  pcm\_fftMag\_psySharpness\_sma\_amean  audspec\_lengthL1norm\_sma\_de\_stddev  pcm\_fftMag\_spectralRollOff50.0\_sma\_amean |

1. Best 3 algorithms

English => MLP, RF, KNN

KNN: acc = 0.66, +-0.03, values: [0.62 0.68 0.66 0.68 0.66 0.67 0.62 0.72 0.63 0.67]

Linear SVM: acc = 0.63, +-0.04, values: [0.6 0.69 0.63 0.59 0.66 0.59 0.64 0.68 0.61 0.63]

RBF SVM: acc = 0.49, +-0.04, values: [0.49 0.48 0.43 0.49 0.43 0.48 0.5 0.59 0.48 0.5 ]

Gaussian Process: acc = 0.37, +-0.10, values: [0.44 0.23 0.44 0.42 0.41 0.43 0.21 0.23 0.44 0.45]

Neural Net: acc = 0.71, +-0.04, values: [0.68 0.73 0.72 0.65 0.74 0.7 0.76 0.73 0.65 0.71]

Naive Bayes: acc = 0.52, +-0.03, values: [0.52 0.58 0.51 0.49 0.49 0.53 0.52 0.5 0.52 0.54]

Random Forest: acc = 0.70, +-0.03, values: [0.68 0.74 0.74 0.64 0.73 0.71 0.71 0.71 0.67 0.71]

Chinese => KNN, RF, MLP

KNN: acc = 0.45, +-0.05, values: [0.51 0.5 0.41 0.44 0.42 0.48 0.51 0.4 0.34 0.46]

Linear SVM: acc = 0.32, +-0.04, values: [0.37 0.26 0.25 0.38 0.34 0.25 0.36 0.33 0.34 0.32]

RBF SVM: acc = 0.26, +-0.04, values: [0.28 0.23 0.29 0.32 0.25 0.17 0.28 0.26 0.22 0.3 ]

Gaussian Process: acc = 0.33, +-0.03, values: [0.3 0.35 0.37 0.33 0.34 0.28 0.28 0.33 0.37 0.36]

Neural Net: acc = 0.42, +-0.03, values: [0.45 0.44 0.45 0.42 0.43 0.4 0.42 0.36 0.44 0.41]

Naive Bayes: acc = 0.29, +-0.03, values: [0.29 0.34 0.27 0.3 0.36 0.27 0.25 0.27 0.26 0.28]

Random Forest: acc = 0.43, +-0.05, values: [0.46 0.46 0.44 0.51 0.43 0.37 0.39 0.34 0.39 0.47]

Turkish => MLP, RF, KNN

KNN: acc = 0.85, +-0.03, values: [0.84 0.82 0.79 0.86 0.87 0.85 0.82 0.9 0.91 0.83]

Linear SVM: acc = 0.82, +-0.02, values: [0.84 0.84 0.8 0.85 0.81 0.82 0.8 0.8 0.87 0.82]

RBF SVM: acc = 0.55, +-0.06, values: [0.58 0.67 0.55 0.45 0.57 0.56 0.47 0.61 0.53 0.51]

Gaussian Process: acc = 0.69, +-0.17, values: [0.88 0.45 0.63 0.85 0.63 0.86 0.82 0.45 0.47 0.83]

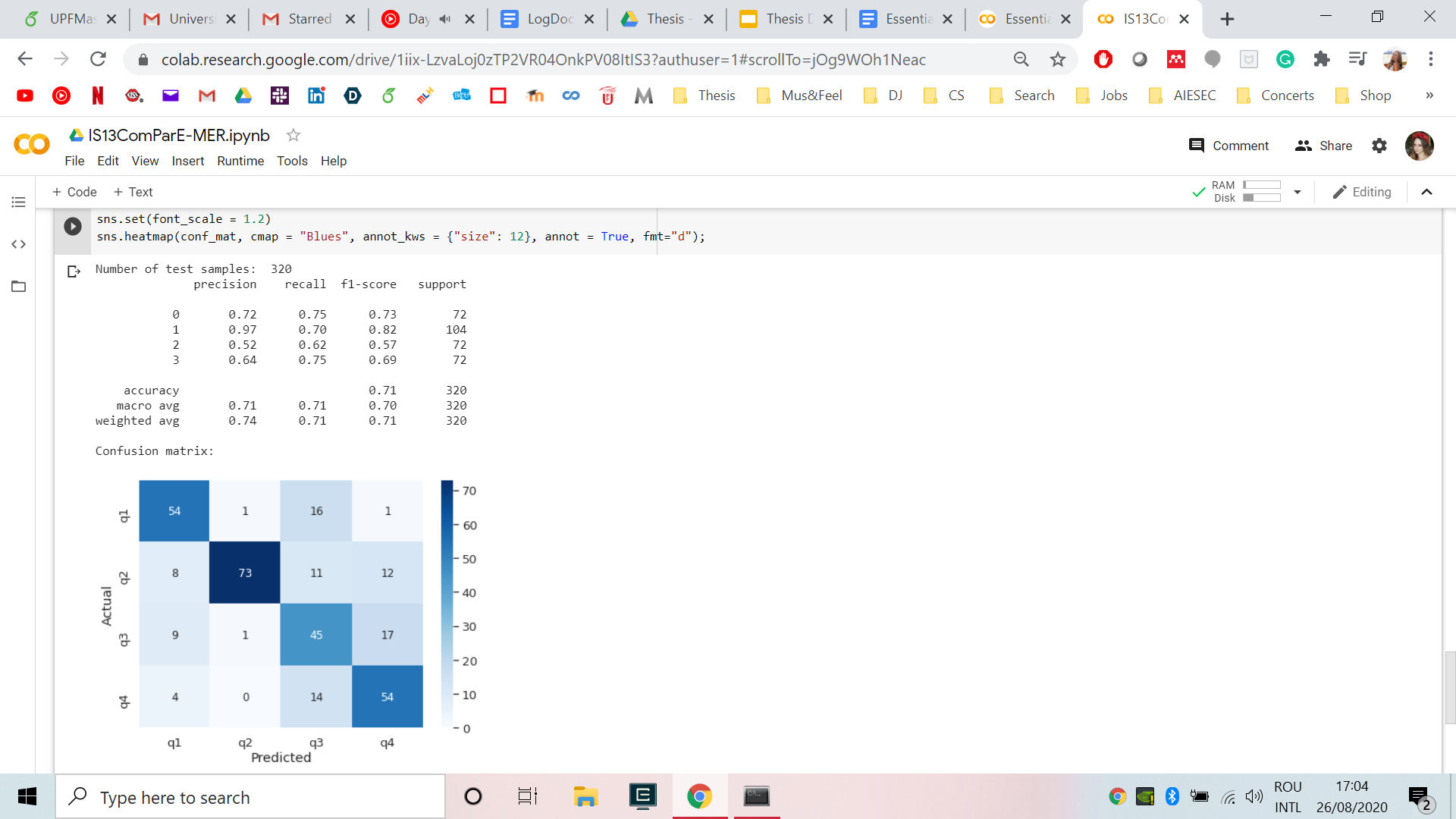
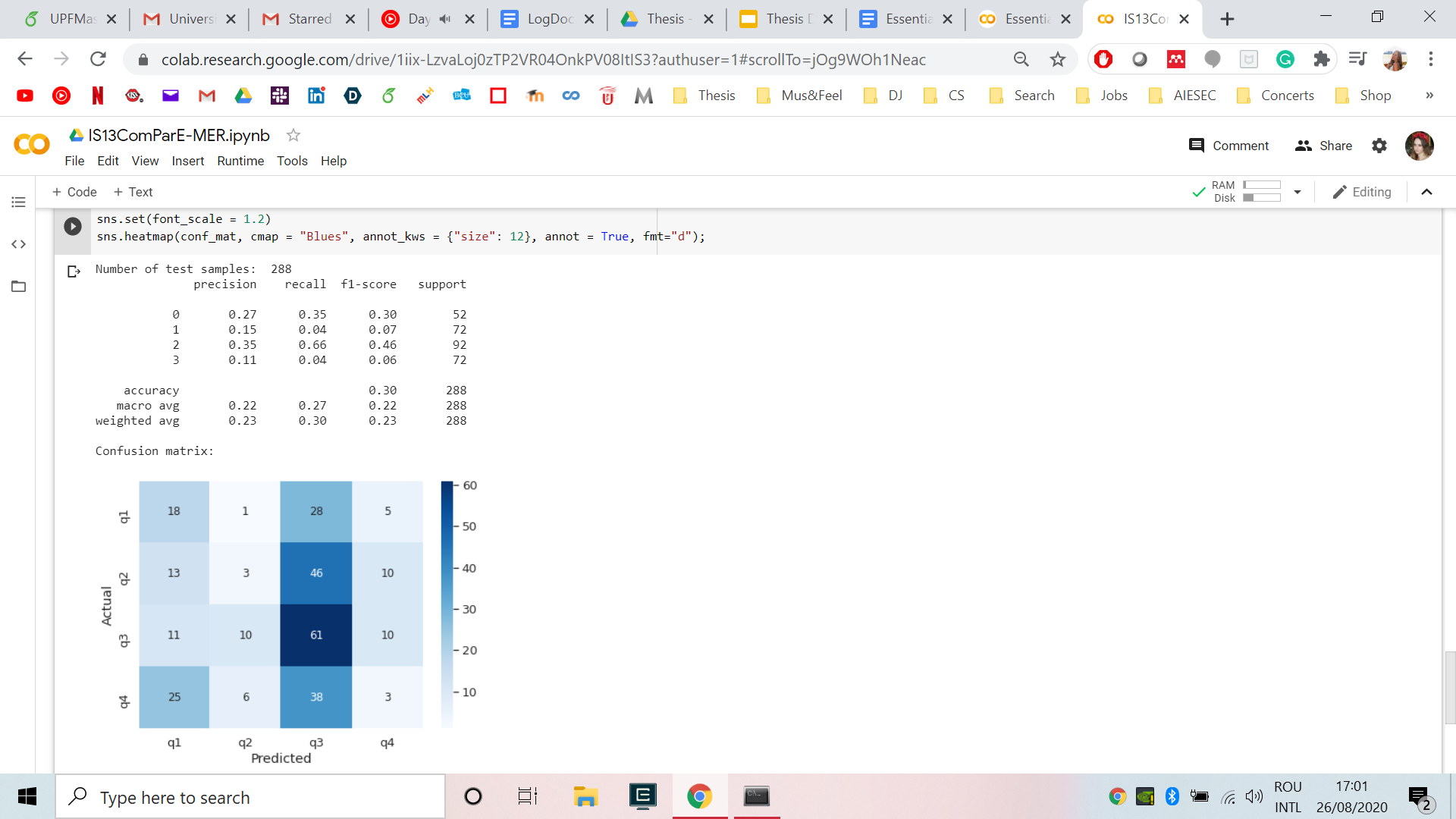
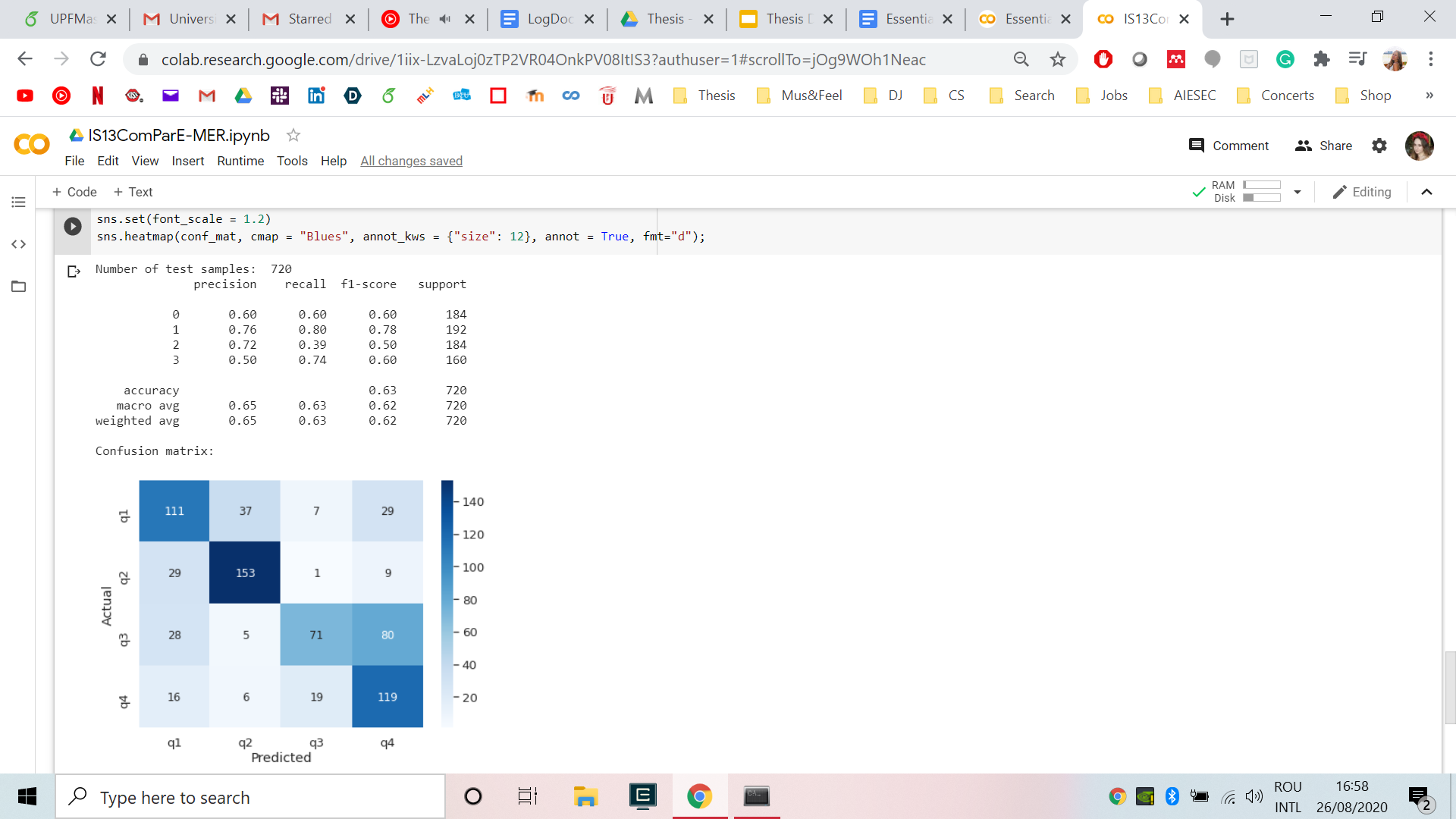
Neural Net: acc = 0.88, +-0.02, values: [0.88 0.92 0.85 0.87 0.88 0.87 0.86 0.86 0.88 0.88]

Naive Bayes: acc = 0.70, +-0.03, values: [0.67 0.7 0.67 0.76 0.66 0.69 0.73 0.7 0.76 0.71]

Random Forest: acc = 0.86, +-0.02, values: [0.87 0.88 0.85 0.91 0.84 0.83 0.86 0.84 0.89 0.85]

1. MLP results

|  |  |  |  |
| --- | --- | --- | --- |
|  | English | Chinese | Turkish |
| Precision | 65% | 23% | 74% |
| Recall | 63% | 30% | 71% |
| F-score | 62% | 23% | 71% |
| Accuracy | 63% | 30% | 71% |



**Cross-dataset**

1. Best 10 features for mixed training

|  |  |
| --- | --- |
| Essentia | IS13ComParE |
| spectral\_complexity.stdev  spectral\_complexity.mean  spectral\_centroid.mean  barkbands\_spread.mean  spectral\_energyband\_high.mean  spectral\_skewness.stdev  melbands\_spread.mean  zerocrossingrate.mean  barkbands\_flatness\_db.mean  silence\_rate\_30dB.stdev | frameTime  logHNR\_sma\_stddev  audspec\_lengthL1norm\_sma\_amean  F0final\_sma\_stddev  F0final\_sma\_amean  pcm\_fftMag\_spectralRollOff75.0\_sma\_amean  pcm\_zcr\_sma\_amean  pcm\_fftMag\_fband1000-4000\_sma\_amean  pcm\_fftMag\_spectralRollOff50.0\_sma\_amean  logHNR\_sma\_amean |

1. Cross-dataset

2.1. Essentia (MLP)

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Train\Test | English | | | | Chinese | | | | Turkish | | | |
| Score(%) | P | R | F | A | P | R | F | A | P | R | F | A |
| English | 60 | 55 | 54 | 55 | 28 | 28 | 23 | 28 | 55 | 47 | 45 | 48 |
| Chinese | 31 | 37 | 29 | 37 | 22 | 20 | 20 | 20 | 23 | 22 | 21 | 22 |
| Turkish | 55 | 52 | 50 | 52 | 19 | 29 | 22 | 29 | 72 | 69 | 70 | 69 |

2.2. IS13ComParE (MLP)

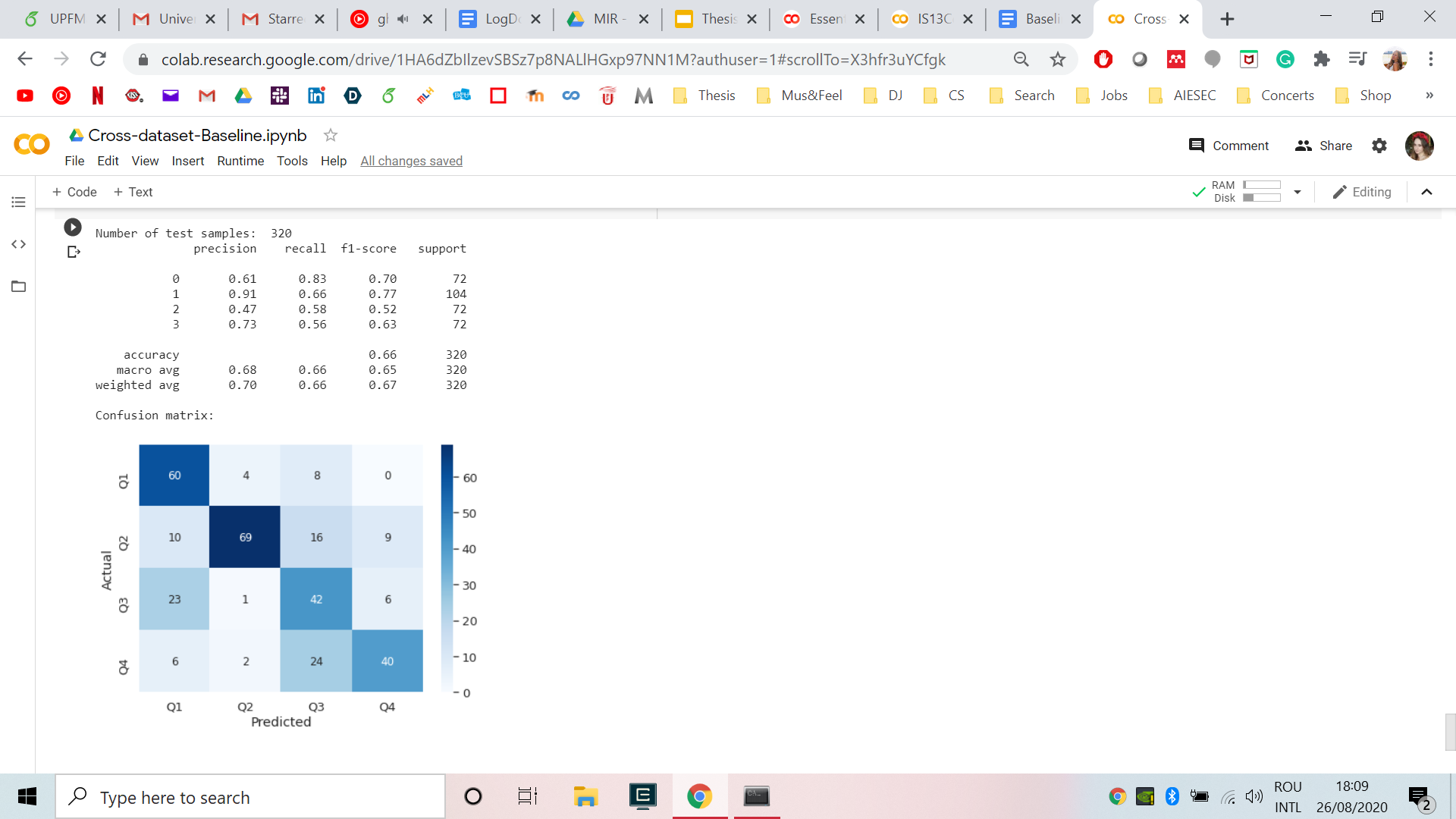
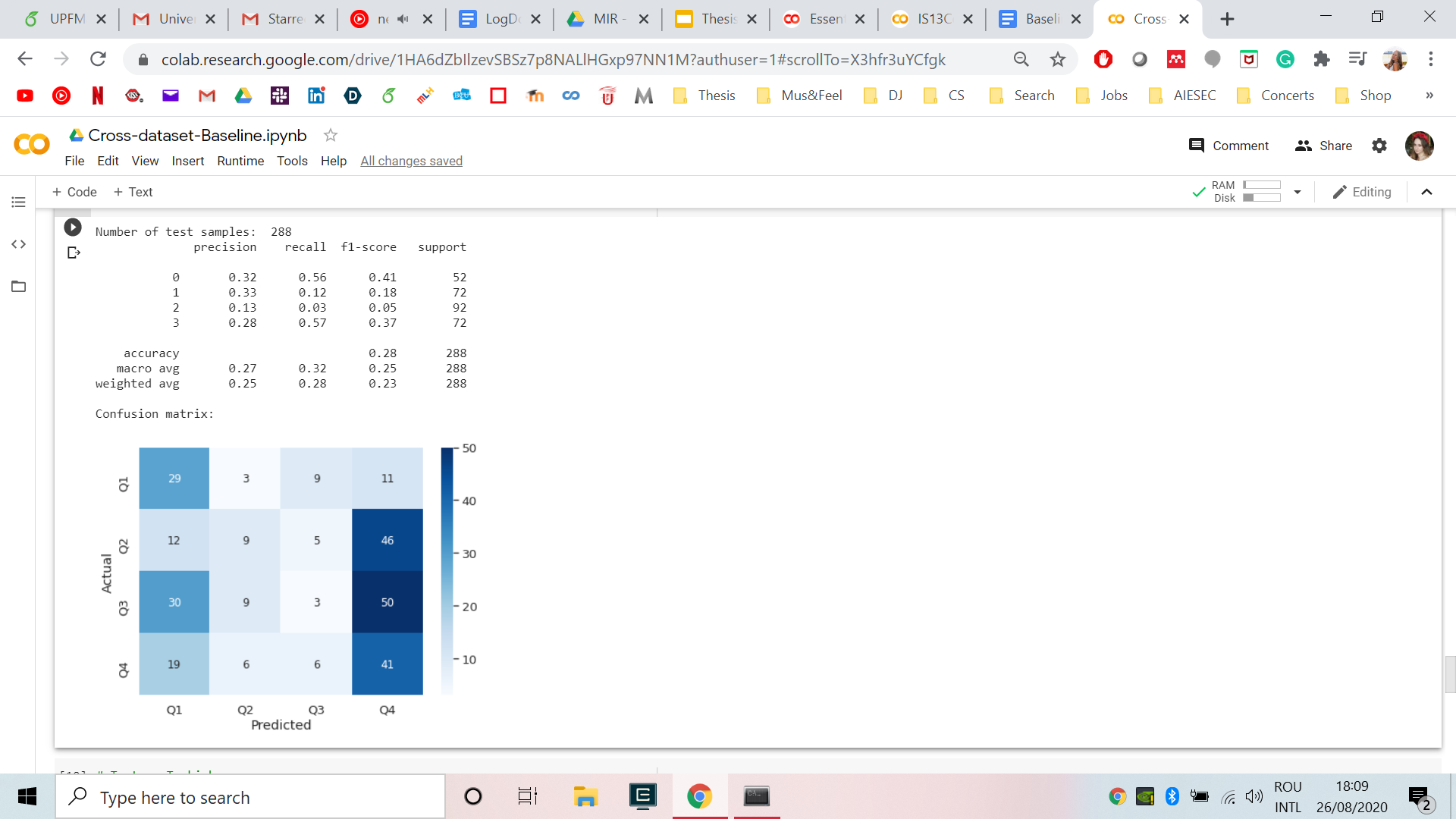
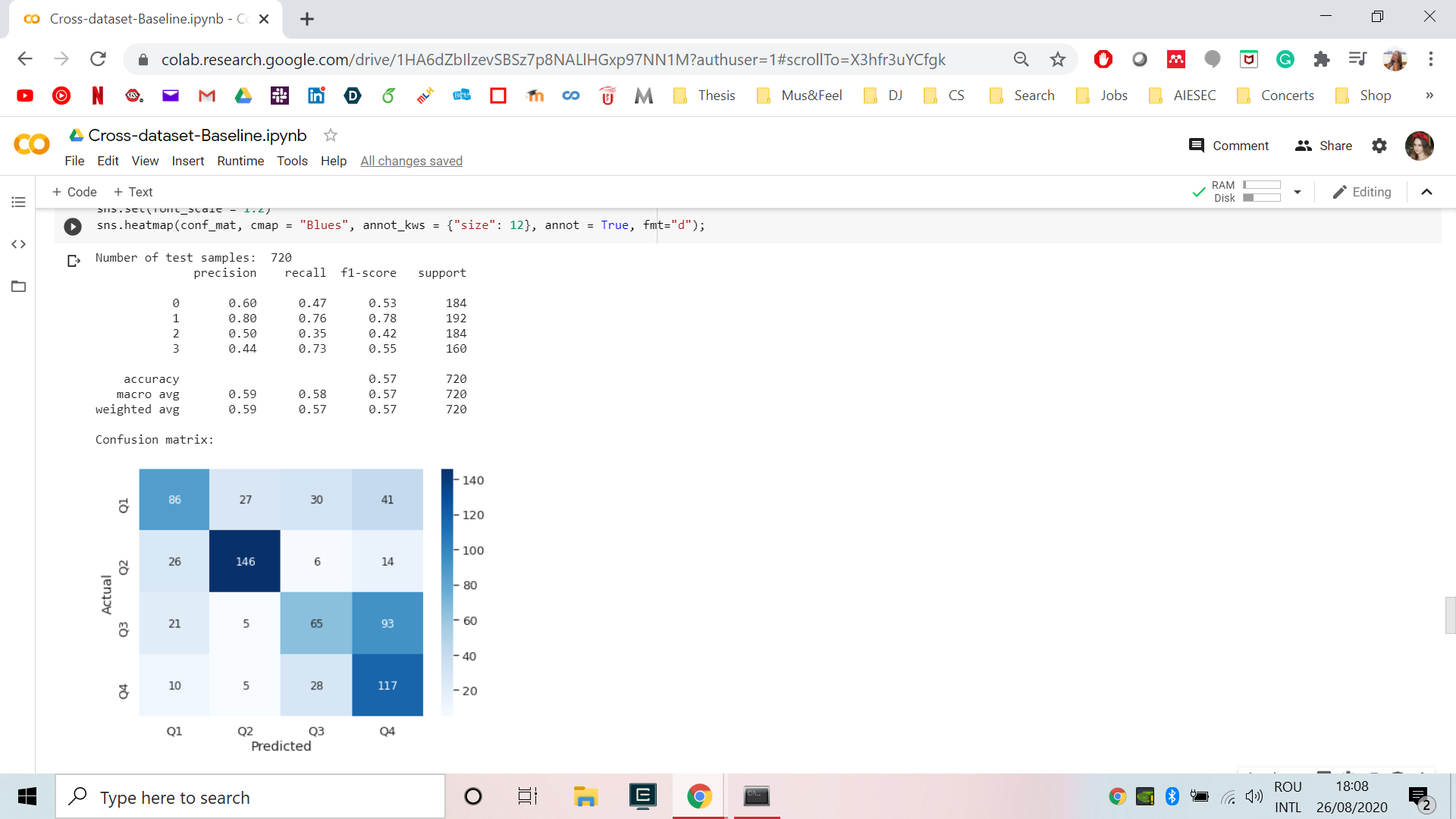
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Train\Test | English | | | | Chinese | | | | Turkish | | | |
| Score(%) | P | R | F | A | P | R | F | A | P | R | F | A |
| English | 59 | 59 | 58 | 59 | 23 | 20 | 20 | 20 | 47 | 46 | 45 | 46 |
| Chinese | 20 | 23 | 20 | 23 | 22 | 23 | 21 | 23 | 26 | 28 | 26 | 28 |
| Turkish | 53 | 45 | 41 | 45 | 23 | 21 | 18 | 21 | 71 | 64 | 64 | 64 |

1. Mixed training (MLP)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Test\Features | Essentia | | | | IS13ComParE | | | |
| Score(%) | P | R | F | A | P | R | F | A |
| English | 59 | 57 | 57 | 57 | 60 | 57 | 56 | 57 |
| Chinese | 25 | 28 | 23 | 28 | 22 | 23 | 22 | 23 |
| Turkish | 70 | 66 | 67 | 66 | 67 | 64 | 64 | 64 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | English | Chinese | Turkish |
| Precision | 60% | 22% | 67% |
| Recall | 57% | 23% | 64% |
| F-score | 56% | 22% | 64% |
| Accuracy | 57% | 23% | 64% |

3.1 Essentia



3.2. IS13ComParE

