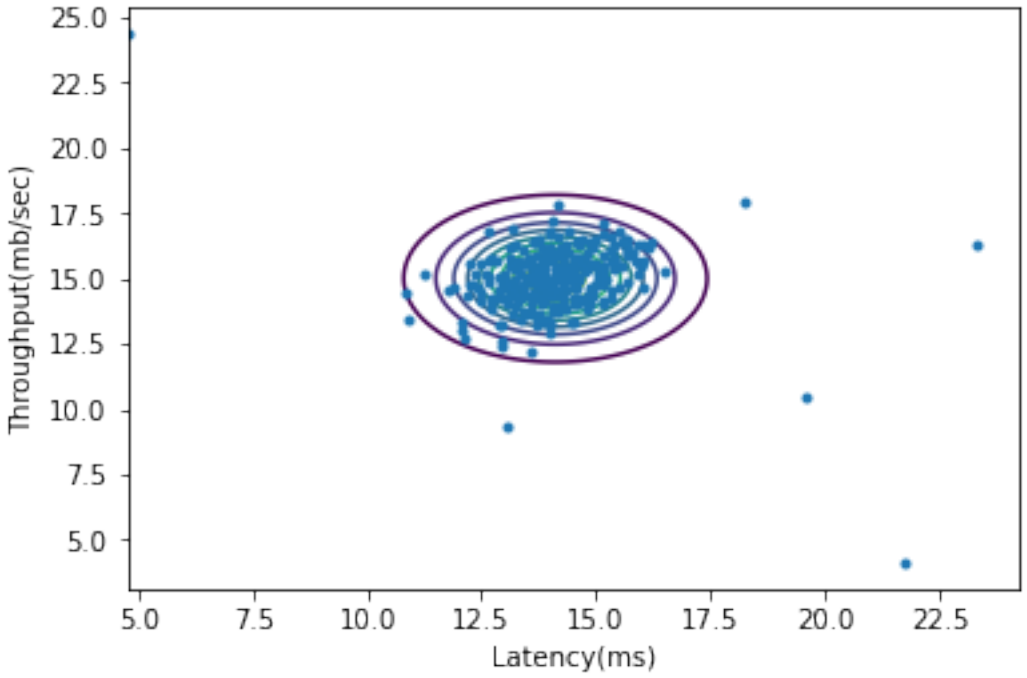
anomaly

February 24, 2021

0.1 Anomaly Detection

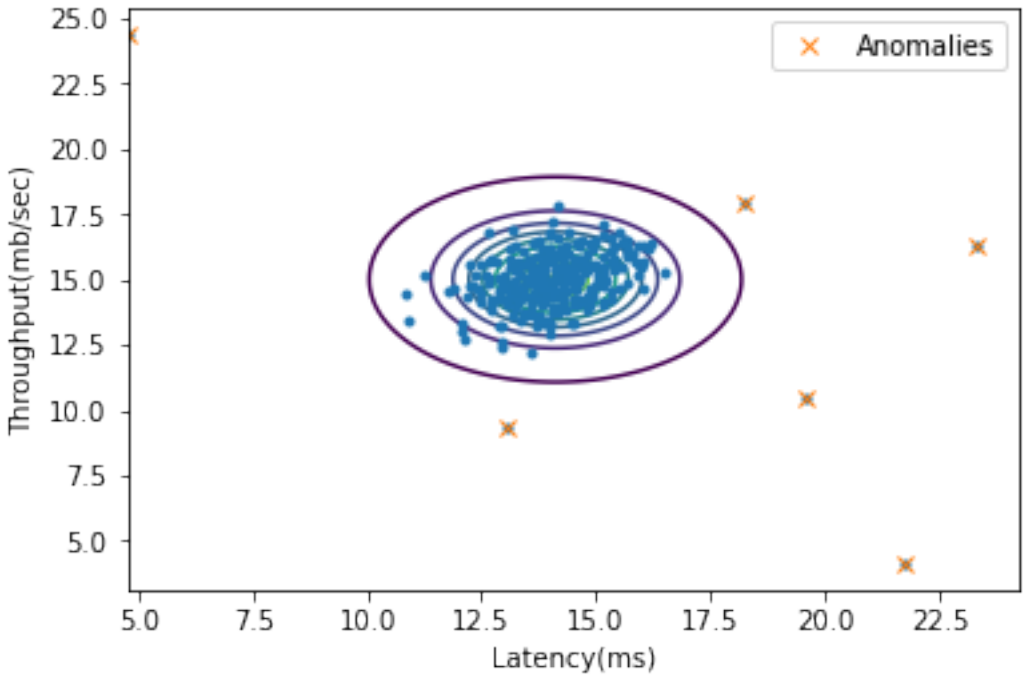
[1]:
```
import ex8_plt
```



The above data will be used for testing an anomaly detection algorithm. Details of the data and benchmarks can be found in Ex8.pdf. This is 2D data for detecting anomalous behavior of server components with m=307 training examples. The data has been fitted with independent Gaussian models for each of the features. The contours of the Gaussian above reflect the fact that most of the data fall in the high probability contours while the anomalous training examples are assigned low probabilities by the fitted model. The following figure shows the anomalies detected with the best threshold that maximises the F1 score on the validation set.

[2]:
```
import ex8_score
```

```
The best threshold obtained is  8.970538225418937e-05
```



Next we consider a much higher dimensional realistic dataset where the behaviors of the servers are characterised by 11 features.

[1]:
```
import ex8_highd
```

```
The threshold obtained for the high dimesional dataset is 1.3730337086886958e-18
Number of anomalies detected equal to 117
```