Supervised Machine Learning: Training

```
data = get from database (full-load into memory, out-of-core)

number_of_epochs = 10
batch_size = 50
number_of_batches = size(data) / batch_size
model = initialize randomly
for i = 1 to number_of_epochs:
    for j = 1 to number_of_batches:
        batch_data = extract batch j from data
        training_with_batch(model, batch_data)

output model
```

Unsupervised Machine Learning: Given the data, find top K outliers

```
for each record R in data:
    score[R] = outlier_score(R, data)
return score

outlier_score(R, data):
score = 0
for each record T in data:
    score += distance(R, T)
return score
```

K-means clustering