

Project Report: Classification

2018.05.27

Tran Dinh Son

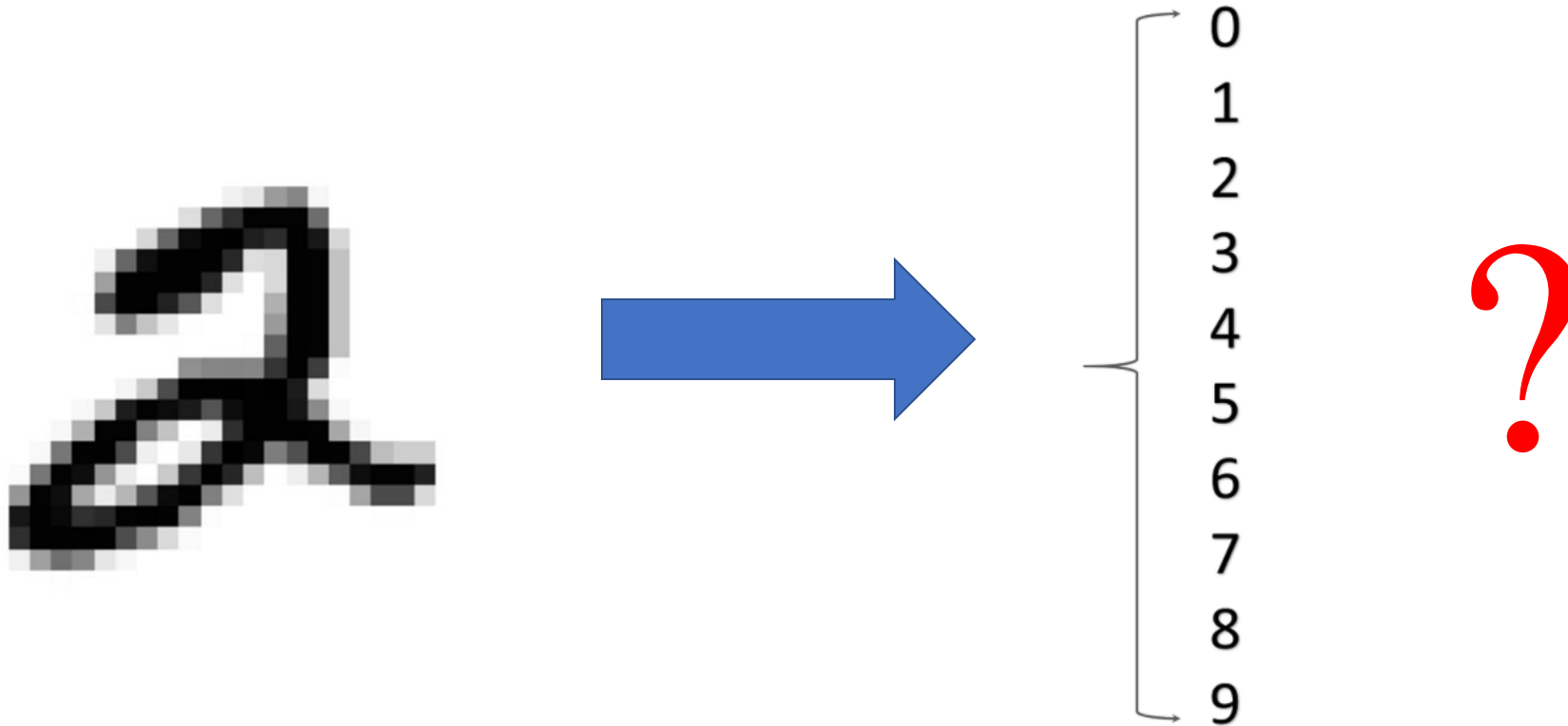
Contents

1. Problem Definition
2. Proposed Method
3. Experimental Results
4. Conclusion

Contents

- 1. Problem Definition**
2. Proposed Method
3. Experimental Results
4. Conclusion

Problem Definition 1



MNIST dataset: **60,000 data** (training set) and **10,000 data (testing set)**

Applications: Human-machine interaction, surveillance, etc.

Dataset repository: <http://yann.lecun.com/exdb/mnist/>

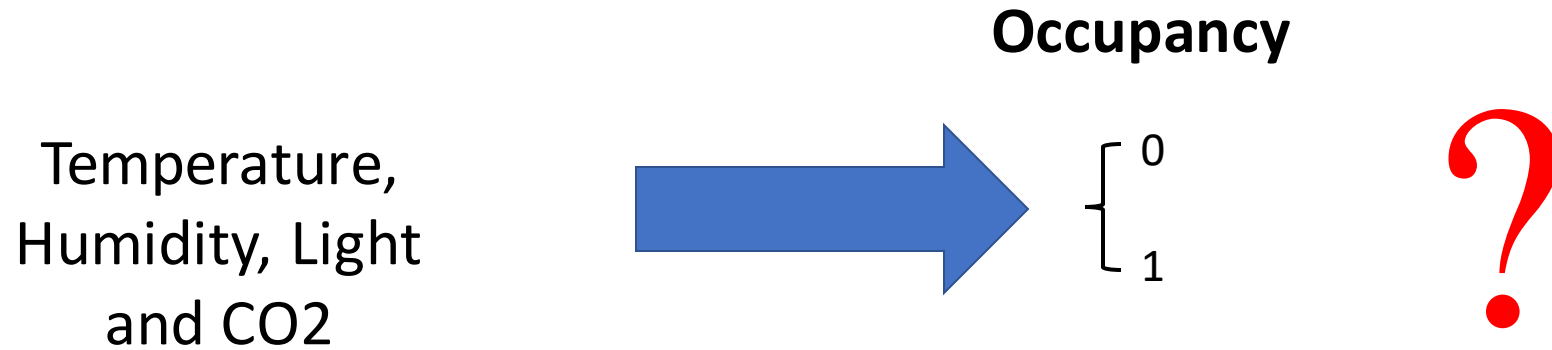
Problem Definition 2



LFW Face dataset: **1,288 data** (70% training set and 30 testing set), **7 class**

Dataset repository: <http://vis-www.cs.umass.edu/lfw/>

Problem Definition 3



Occupancy Detection Dataset:

Dataset repository: <https://archive.ics.uci.edu/ml/datasets/Occupancy+Detection+#>

Contents

1. Problem Definition
- 2. Proposed Method**
3. Experimental Results
4. Conclusion

Proposed Method

- Multilayer perceptron (MLP)
- Decision Tree
- Support vector machine (SVM)

Contents

1. Problem Definition
2. Proposed Method
- 3. Experimental Results**
4. Conclusion

Comparison on MNIST

Methods	Accuracy
MLP	96.98%
Decision Tree	88%
SVM	94.03%

The materials for this section is on
<https://github.com/trandinhs086/Data-Mining-class>

Comparison on Occupancy Detection Dataset

Methods	Accuracy
MLP	98.83%
Decision Tree	98.88%
SVM	90.95%

The materials for this section is on
<https://github.com/trandinhsen3086/Data-Mining-class>

Comparison on LFW Face Dataset

Methods	Accuracy
MLP	81.05%
Decision Tree	77.20%
SVM	84.47%

The materials for this section is on
<https://github.com/trandinhsen3086/Data-Mining-class>

Contents

1. Problem Definition
2. Proposed Method
3. Experimental Results
- 4. Conclusion**

Conclusion

- Depend on feature of dataset, each dataset has its own advantage. Three cases above prove that.
- MLP seems to show better results than others in all three cases

THANK YOU!