A comparison of classifiers for oil spill detection

Andy
Thomas
Soheil

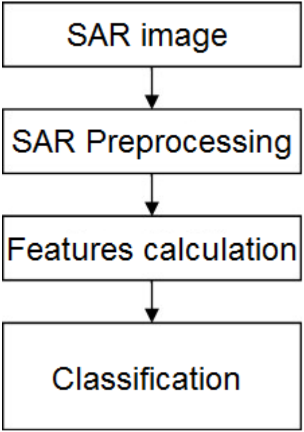
Delft University of Technology

30th October 2014

Overview

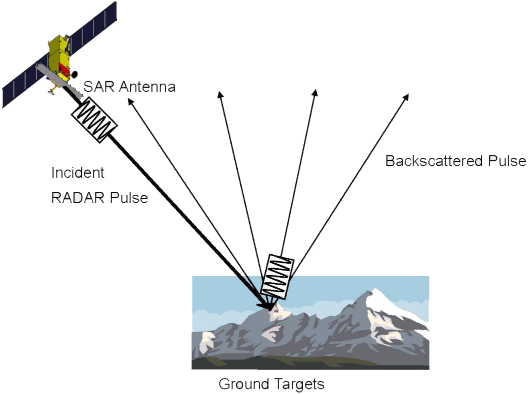
What are Oil Spills?

General oil spill detection approach

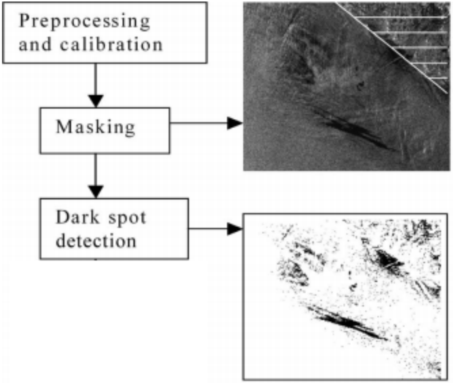


Synthetic Aperture Radar Image

- Radar?
- Synthetic Aperture?
- Advantages
- Problems



SAR Preprocessing

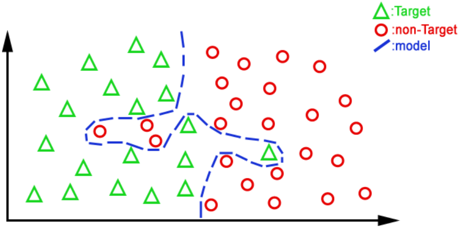


Features

- What are features?
- Types of features
  - Geometrical
  - Physical
  - Texture
  - Contextual

Feature Selection

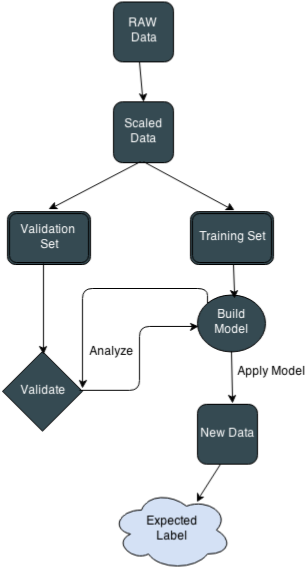
- Choosing features
- Over fitting
- Curse of dimensionality



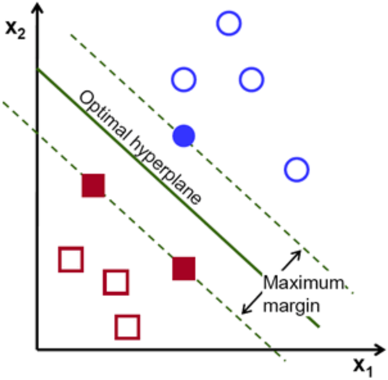
Typical problems

- Lookalikes
- Imbalanced dataset
- Data is scarce
- Gathering contextual features

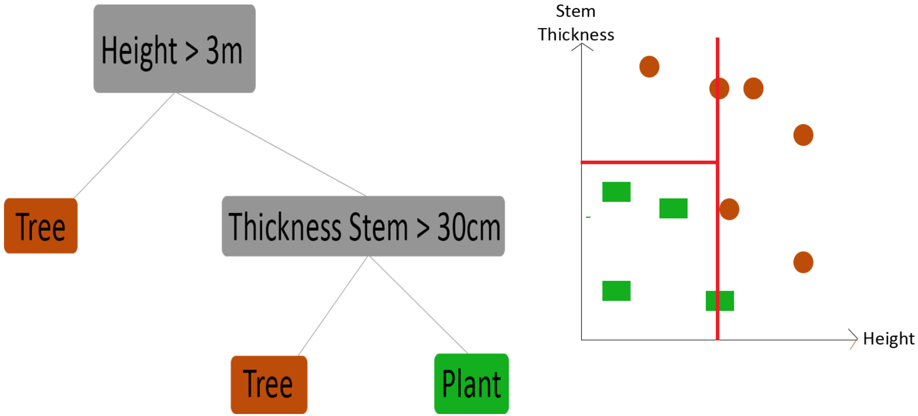
Supervised learning



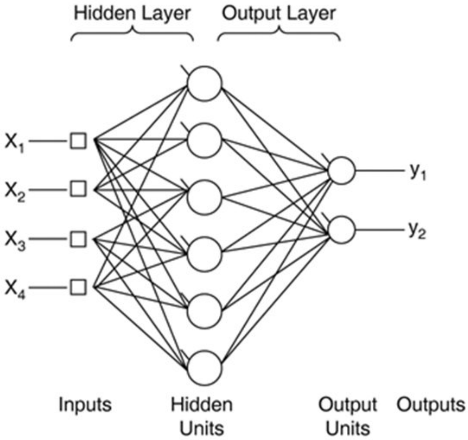
Support vector machines



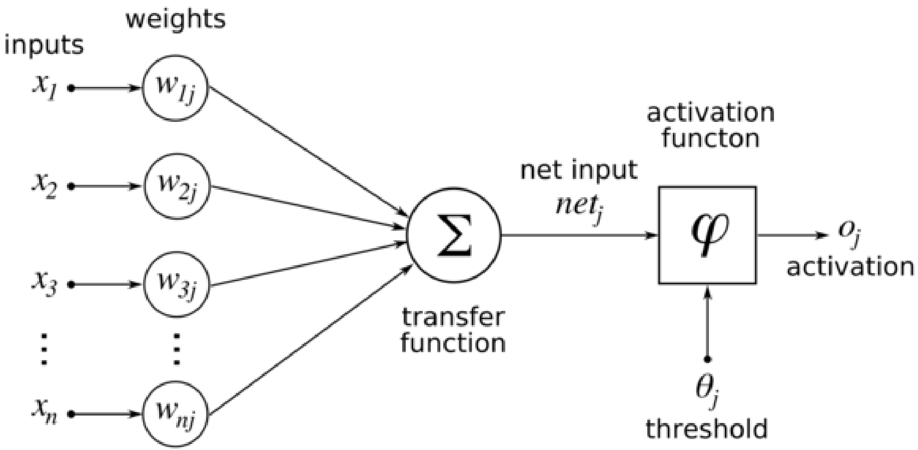
Decision Tree



Multilayer perceptrons

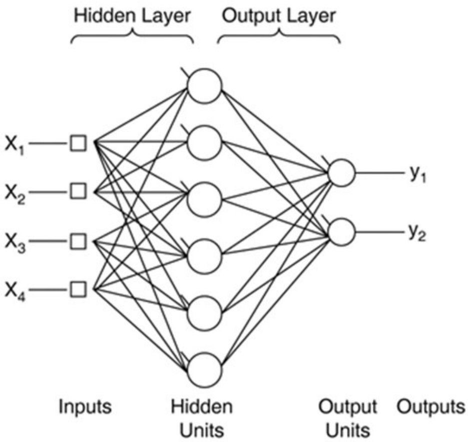


Perceptron



weights
inputs
x_1
w_{1j}
x_2
w_{2j}
x_3
w_{3j}
x_n
w_{nj}
Σ
transfer function
net input
net_j
activation functon
φ
o_j
activation
θ_j
threshold

Multilayer perceptrons



Research goals

- Accuracy comparison
- Classifier characteristics vs field specific problems

Difficulties researching

- SVM was hardly used
- Not all details were specified
- Use of different datasets (unverified samples)
- Use of different features
- Free lunch theorem

Fail

We did not manage to compare our three classifiers

The best classifier?

- Different accuracy in different situations
- SVM and MLP for large datasets
- SVM and MLP for non-linear cases
- DT for small datasets
- DT easiest to interpret

Recommendations

- More research on SVM
- More research on Random Forest
- Shared database
- Bagging
- Image fusion

Summary

- Oil spill detection
- Classifiers
- Research
- Difficulties
- Recommendations

Study	Data Type	Preprocessing	# Features	Formations(lookalike & oilspills)	Results
Oil spills[59]	ERS-2 SAR, 24 high-res images 8-bit	transformation, Filtering, data normalization	10	90 & 69	MLP(10:51:2) accuracy: 86.67% lookalike acc. 91.18% oil spills acc.
Oil spills[17]	ERS SAR, 600 low-res images	Resampling,Radiometric range correction, georeference	11	68 & 71	MLP(11-8-4-1) accuracy: 90% lookalike acc. 82% oil spills acc.
Oil spills[62]	ERS-2 SAR, 24 high-res images	-	10	90 & 69	MLP(10-51-2) accuracy: 84.4% lookalike acc. 85.3% oil spills acc.
Oil spills[60]	ERS SAR, 12 high-res images	8-bit transformation, filtering	-	-	MLP(4-2-1) accuracy: 96.46% overall acc.
Oil spills[18]	ERS SAR, 70 images	-	12	78 & 111	MLP(12-8-8-1) 0.227 root mean square error(rmse)
Oil spills[66]	RADARSAT-1, 93 images	log-transformation, standardization	15	94 & 98	MLP 75.93% overall acc. SVM 79.63% overall acc. DT(Bundling) 90.74% overall acc.
Oil spills[37]	Envisat, 47 images	-	9	155 & 80	MLP(9:11:2) 96.3% lookalike acc. 92.9% oil spill acc.
					DT 92.6% lookalike acc. 92.9% oil spill acc.
Oil spills[56]	ERS-1 SAR, 59 low-res images	-	9	2471 & 42	DT 96% lookalike acc. 86% oil spill acc
Oil spills[61]	ERS-2 SAR, 24 high-res images	-	9	90 & 69	DT forest 84.4%
Hydro-acoustics [50]	Sonar	Echoview	15	-	SVM accuracy: 89.5%, DT accuracy: 86.8%
Land coverage1986 [2]	LandSat SAR	'Using data miner'	11	-	SVM max accuracy: 90.53%, DT accuracy: 93.48%
Land coverage2001 [2]	LandSat SAR	'Using data miner'	10	-	SVM max accuracy: 93.67%, DT accuracy: 94.07%
ECG arrhythmias[40]	MIT-BIH arrhythmia database	-	10	-	accuracy SVM 99% , MLP 98.22%
Remote Sensing[68]	Satimage	feature extraction	26	-	accuracy for SVM 93.16% ,and for MLP 96.98%
signature recognition[19]	user signature data	feature extraction	2	-	Recognition rate SVM 66.5 , MLP 71.2
wind speed prediction[41]	daily wind speed data	-	high dimensional feature	-	MSE on testing data for the MLP is 0.0090 while it is 0.0078 for the SVM
Hashimoto's disease[44]	66 Thyroid ultrasound images	normalization	59	54 healthy and 85 sick	MLP(6-8-1): 89.4% sick class 61.1% healthy class. DT: 89.4% sick class 94.4% healthy class.

In case you need our credentials
A.S.Y.Chiu 1519360
T.P.van Helden 4106725
S.S.Jahanshahi 4127617