ML EXAMS 2021/01/18

EX. 1

1.

is a regression problem and we have a ~~fome~~ target function $f: X \to Y$ with $X = \{x_1, x_2, x_3\}$ and $Y = \{4.6, 21, 10\}$ and dataset $D = \{(x_n, f_n)\}_{n=1}^{N}$

for solve this task I will use regression model define in this way $y(x,w) = \sum_{n=1}^{N} w_n^T \phi_n(x)$

and the goal of this function is minimize this:

$$\arg\min E_D(w) = \arg\min \frac{1}{2} \sum_{n=1}^{3} (f_n - w^T\phi(x_n))^2$$

2 because we want find $w^* = \arg\min_w E_D(w)$ we can use sequential algorithm

$$\hat{w} \leftarrow \hat{w} + \eta [f_n - w^T\phi(x_n)]\phi(x_n)$$

EX. 2

1

yes, data is linearly separable infact exist an surface that divide our instances space into two regions such that different classified instances are separeted

2.

due the fact that in this examples there are some outliers and from the images we can see the result is sensible to outliers the model used is LEAST SQUARES

3.

SVM to due the fact aim to maximum ~~with the~~ margin for better accuracy for this reason is robust to outliers.

EX. 3

1

is binary classification task $f: X \to \{+, 1\}$ if we consider a dataset very unbalanced for example with 95% of samples positive an hypothesis $h_1(x)$ that respond only + for every samples in input will have an high accuracy respect to anothere hypothoses $h_2(x)$ that respond + or -

2. a more reliable metrics are

precision = $\frac{TP}{TP + FP}$ recall = $\frac{TP}{TP + FN}$ F1-score = $\frac{2(\text{precision} \cdot \text{recall})}{\text{precision} + \text{recall}}$

EX 4

1.

when the input vector appear in the form of inner product $x^T x'$
we can update it with the kernel function $k(x, x')$

2.

we can consider SVM for classification we know that $w^* = \sum_{n=1}^{N} a_n^* t_n x_n$ and

$w^* = w_0^* + \sum_{x_j \in SV} a_j t_j x_j^T x_j = 0$ with a^* LAGRANGE MULTIPLIERS and

$SV = \{ x_k \in D : t_k y(x_k) = 1 \}$

now we kernelize this method:

$$y(x, w^*) = \text{sign}\left(w_0 + \sum_{n=1}^{N} a_n^* x^T x \right)$$

$$y(x, w^*) = \text{sign}\left(w_0 + \sum_{n=1}^{N} a_n^* k(x^T x) \right)$$

with a polynomial
kernel function

$$y(x, w^*) = \text{sign}\left(w_0 + \sum_{n=1}^{N} a_n^* (\beta x^T x + \gamma)^d \right)$$ with

$$w_0 = \frac{1}{|SV|} \sum_{x_i \in SV} \left(t_i - \sum_{x_j \in SV} a_j^* t_j k(x_i, x_j) \right)$$

EX. 5

1.

outputs units as SOFTMAX activation functions:

$$y_i = \text{softmax}(\alpha^i) = \frac{\exp(\alpha^{(i)})}{\sum_j \exp(\alpha_j)}$$

2. loss function: categorical loss function:

$$J(\theta) = E_{x,t \sim D}\left[-\ln(\text{softmax}(\alpha^{(i)})) \right]$$

with $\alpha^{(i)} = w_i^T h + b$

outputs units saturate only when there are minimal errors

EX 6

1\.

try to solve the problem of dimensionality reduction
feature extraction and compression visualization

2\.

3	3	3
E	3	m

we have 36x36x3 RGB images
the dimensionality reduction of dataset is
36x36x3 but the intrinsic paramete
are 3 2 for translation one for
rotation so with dimensionality reduction

we try to $\mathbb{R}^{w \times h \times d} \rightarrow \mathbb{R}^3$

ML EXAMS 16/01/2023

EX. 1

1.1

if $C = C_1$ and $B = b_1$ then NO (1)
if $C = C_1$ and $B = b_2$ then YES (2)

if $C = C_2$ and $A = a_1$ then YES (3)
if $C = C_2$ and $A = a_2$ and $B = b_1$ then YES (4)
if $C = C_2$ and $A = a_2$ and $B = b_2$ then YES (5)
if $C = C_2$ and $A = a_3$ then NO (6)

if $C = C_3$ then NO (7)

1.2

S_1 is consistent because of (1)
S_2 is consistent because of (4)
S_3 is consistent because of (7)
S_4 is NOT consistent because of (2)

1.3

accuracy $\frac{3}{4} = \frac{\text{number of consistent instances}}{\text{number of all instances}} = 0.75$