Artificial Intelligence III: Artificial Intelligence and Deep Learning

Ch05 –Ensemble

Tutorial Answer

1. Given three base classifiers in a ensemble and 5 testing samples as below:

Testing Sample	f_1			f_2			f_3			True Output
	ω_1	ω_2	ω_3	ω_1	ω_2	ω_3	ω_1	ω_2	ω_3	
x_1	0.7	0.2	0.3	0.5	0.2	0.6	0.7	0.4	0.1	ω_1
x_2	0.4	0.6	0.1	0.2	0.2	0.9	0.1	0.4	0.2	ω_3
x_3	0.2	0.2	0.5	0.5	0.2	0.4	0.7	0.5	0.5	ω_1
x_4	0.5	0.3	0.6	0.4	0.5	0.1	0.1	0.1	0.4	ω_3
x_5	0.2	0.9	0.8	0.0	0.1	0.2	0.2	0.6	0.7	ω_2

a) What is the testing accuracy of each base classifier?

Answer:

Testing Sample	f_1	f_2	f_3	True Output
x_1	ω_1	ω_3	ω_1	ω_1
x_2	ω_2	ω_3	ω_2	ω_3
x_3	ω_3	ω_1	ω_1	ω_1
x_4	ω_3	ω_2	ω_3	ω_3
x_5	ω_2	ω_3	ω_3	ω_2

Testing accuracy of f_1 = 60%
Testing accuracy of f_2 = 40%
Testing accuracy of f_3 = 60%

b) If MAJORITY vote is used as the fusion method, which class does the samples (x_i, i = 1..5) belongs to? What is the testing accuracy of this ensemble?

Answer:

Testing Sample	f_1	f_2	f_3	ensemble (MVote)	True Output
x_1	ω_1	ω_3	ω_1	ω_1	ω_1
x_2	ω_2	ω_3	ω_2	ω_2	ω_3
x_3	ω_3	ω_1	ω_1	ω_1	ω_1
x_4	ω_3	ω_2	ω_3	ω_3	ω_3
x_5	ω_2	ω_3	ω_3	ω_3	ω_2

Testing accuracy of ensemble using majority vote = 60%

c) If MAX is used as the fusion method, which class does the samples (x_i, i = 1..5) belongs to? What is the testing accuracy of this ensemble?

Answer:

	f_1			f_2			f_3			*ensemble*			ensemble Output	True Output
	ω_1	ω_2	ω_3	ω_1	ω_2	ω_3	ω_1	ω_2	ω_3	ω_1	ω_2	ω_3		
x_1	0.7	0.2	0.3	0.5	0.2	0.6	0.7	0.4	0.1	0.7	0.4	0.6	ω_1	ω_1
x_2	0.4	0.6	0.1	0.2	0.2	0.9	0.1	0.4	0.2	0.4	0.6	0.9	ω_3	ω_3
x_3	0.2	0.2	0.5	0.5	0.2	0.4	0.7	0.5	0.5	0.7	0.5	0.5	ω_1	ω_1
x_4	0.5	0.3	0.6	0.4	0.5	0.1	0.1	0.1	0.4	0.5	0.5	0.6	ω_3	ω_3
x_5	0.2	0.9	0.8	0.0	0.1	0.2	0.2	0.6	0.7	0.2	0.9	0.8	ω_2	ω_2

Testing accuracy of ensemble using MAX = 100%

d) Assume that the training accuracy of classifier 1, 2 and 3 are 95%, 85% and 95%. By using WEIGHTED AVERAGE as the fusion method, which class does the samples (x_i, i = 1..5) belongs to? The weight should be determined by using this formula:

$$w_i = \frac{err_i}{\sum_{j=1}^{L} err_j}$$, where err_i is the training accuracy of classifier i

What is the testing accuracy of this ensemble?

Answer:

$w_1 = 0.3455$

$w_2 = 0.3091$

$w_3 = 0.3455$

x_1 is class ω_1

0.7*0.3455+0.5*0.3091+0.7*0.3455=0.6382
0.2*0.3455+0.2*0.3091+0.4*0.3455=0.2691
0.3*0.3455+0.6*0.3091+0.1*0.3455=0.3237

x_2 is class ω_2

0.4*0.3455+0.2*0.3091+0.1*0.3455=0.2346
0.6*0.3455+0.2*0.3091+0.4*0.3455=0.4073
0.1*0.3455+0.9*0.3091+0.2*0.3455=0.3818

x_3 is class ω_3

0.2*0.3455+0.5*0.3091+0.7*0.3455=0.4655
0.2*0.3455+0.2*0.3091+0.5*0.3455=0.3037
0.5*0.3455+0.4*0.3091+0.5*0.3455=0.4691

x_4 is class ω_3

0.5*0.3455+0.4*0.3091+0.1*0.3455=0.3309
0.3*0.3455+0.5*0.3091+0.1*0.3455=0.2927
0.6*0.3455+0.1*0.3091+0.4*0.3455=0.3764

x_5 is class ω_3

0.2*0.3455+0.0*0.3091+0.2*0.3455=0.1382
0.9*0.3455+0.1*0.3091+0.6*0.3455=0.5492
0.8*0.3455+0.2*0.3091+0.7*0.3455=0.5801

Testing accuracy of ensemble using weight average = 40%

e) Calculate the diversity for the ensemble by using the following measurements:
   i) Disagreement Measure
   ii) Double Fault Measure
   iii) Correlation coefficient

Answer:

Testing Sample	$f_1 f_2$	$f_2 f_3$	$f_1 f_3$
Dis	5/5	3/5	2/5
DFM	0/5	1/5	1/5

i) Disagreement Measure = 2/3

ii) Double Fault Measure = 2/5

iii) Correlation coefficient

$f_1 f_2 : \omega_1$	0.4892	$f_1 f_3 : \omega_1$	0.1506	$f_2 f_3 : \omega_1$	0.6778
$f_1 f_2 : \omega_2$	-0.4757	$f_1 f_3 : \omega_2$	0.4820	$f_2 f_3 : \omega_2$	-0.9692
$f_1 f_2 : \omega_3$	-0.9284	$f_1 f_3 : \omega_3$	0.8759	$f_2 f_3 : \omega_3$	-0.7048

Mean of all those values is the answer.

f) Do you think adding more base classifiers can increase the performance of the ensemble?

Answer:

The performance of the ensemble can improve only if added base classifiers contribute independently among others.