Maulana Azad National Urdu University

Master of Technology (M.Tech) : II Semester Examination, August 2021

Paper : MTCS211PCT : Machine Learning

Time : 3 hrs **Marks : 70**

ہدایات:

یہ پرچہ سُوالات دو حصوں پر مشتمل ہے: حصہ اول اور حصہ دوم۔ ہر جواب کے لیے لفظوں کی تعداد اشارۃً ہے۔ تمام حصوں سے سوالوں کا جواب دینا لازمی ہے۔

1. حصہ اول میں 10 سوالات ہیں، اس میں سے طالب علم کو کوئی 08 سوالوں کے جواب دینے ہیں۔ ہر سوال کا جواب تقریباً سو (100) لفظوں پر مشتمل ہے ہر سوال کے لیے 05 نمبرات مختص ہیں۔ **(8 x 5 = 40 Marks)**
2. حصہ دوم میں 05 سوالات ہیں۔ اس میں سے طالب علم کو کوئی 03 سوال کا جواب دینا ہے۔ سوال کا جواب تقریباً ڈھائی سو (250) لفظوں پر مشتمل ہے۔ سوال کے لیے 10 نمبرات مختص ہیں۔ **(10x3 = 30 Marks)**

حصہ اول

1. Overfitting کیا ہے؟ ان تمام طریقوں پر Discussion کریں جو Machine Learning Models میں Overfitting ہونے سے روکنے کے لیے استعمال کیے جا سکتے ہیں؟

2. Transactions کے درج ذیل Set پر غور کریں۔ ID3 Algorithm استعمال کریں اور ان Transactions سے ایک Decision Tree کو Construct کریں اور اس کی وضاحت کریں۔

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125 K	No
2	No	Married	100 K	No
3	No	Single	70 K	No
4	Yes	Married	120 K	No
5	No	Divorced	95 K	Yes
6	No	Married	60 K	No
7	Yes	Divoced	220 K	No
8	No	Single	85 K	Yes
9	No	Married	75 K	No
10	No	Single	90 K	Yes

3. Bayes Theorem کیا ہے؟ Maximum Posterior Hypothesis کی وضاحت کریں۔ Decision Tree Induction Method کے Comparison میں Bayes Theorem کے دو فوائد لکھیں۔

4. Support Vector Machine کی وضاحت کریں۔ ANN کے Comparison میں SVM کے فوائد کیا ہیں؟ ANN کے Overffiting Problem کو SVM کیسے Overcome کرتا ہے؟

5. درج ذیل Categories کے لیے Filter Feature Selection Methods کے لیے استعمال کیے جانے والے Statistical Measures کی کم از کم ایک مثال لکھیں۔ Formula بھی لکھیں۔

a. Numerical, Input, Numerical Output
b. Numerical Input, Categorical Output
c. Categorical Input, Numerical Output
d. Categorical Input, Categorical Output

6. K-Nearest Neighbour Algorithm کی تفصیلی وضاحت کریں۔ Euclidean Distance کو Compute کرنے کے لیے Formula لکھیں۔
7. MLP کا Architecture ڈرا کریں اور Back-propagation Model کے ذریعے Prediction کا طریقہ بیان کریں۔
8. دو Machine learning Algorithm کا Comparison کرنے کے طریقوں پر Discuss کریں۔
9. Learning کی تین اہم اقسام کا ذکر کریں۔ کس Category کو دوسری Categories پر ترجیح دی جاتی ہے اور کیوں؟ کسی Learning Method کے Performance کو کیسے Evaluate کیا جا سکتا ہے؟
10. Classification اور Prediction Problems کے ہر Category کے لیے استعمال ہونے والے کسی دو Metrics کا Formula لکھیں۔

حصہ دوم

11. Ada Boost الگورتھم کے Steps لکھیں۔ مثال کے ساتھ وضاحت کریں۔ Voting Function کی وضاحت کریں اور Bagging Algorithm کے Steps لکھیں۔
12. ذیل میں دیے گئے Training Example پر غور کریں جو MRI Scans سے Malignant Tumors تلاش کرتا ہے۔ MRI Scans کے Classification کے لیے Logistic Regression کو Apply کریں۔
13. مناسب مثالوں کے ساتھ Generative Training اور Discriminative Training کے درمیان فرق کریں۔
14. Classification Algorithms کو Assess اور Compare کرنے کے لیے کسی بھی دو طریقوں کو تفصیل سے Discuss کریں۔ Re-sampling اور Cross - Validation میں فرق کریں۔
15. درج ذیل Database پر غور کریں اور نیچے دیے گئے Attributes کے Information Gain کو Compute کریں۔ Apply کیے جانے والا Formula لکھیں۔

a. Age b. Income c. Student_category

age	income	student	credit_rating	buys_computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31..40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31...40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31...40	medium	no	excellent	yes
31....40	high	yes	fair	yes
>40	medium	no	excellent	no

☆☆☆