

Maulana Azad National Urdu University

Ph.D. (Computer Science) Course Work Examination - February 2021

Paper - PHCS105DST : Machine Learning (Backlog)

پرچہ: مشین لرننگ

Time : 3 hrs

Marks : 70

ہدایات:

یہ پرچہ سوالات تین حصوں پر مشتمل ہے: حصہ اول، حصہ دوم، حصہ سوم۔ ہر جواب کے لئے لفظوں کی تعداد اشارہ ہے۔ تمام حصوں سے سوالوں کا جواب دینا لازمی ہے۔

1. حصہ اول میں 10 لازمی سوالات ہیں جو کہ معروضی سوالات/خالی جگہ پُر کرنا/مختصر جواب والے سوالات ہیں۔ ہر سوال کا جواب لازمی ہے۔
ہر سوال کے لیے 1 نمبر مختص ہے۔
(10 x 1 = 10 Marks)

2. حصہ دوم میں آٹھ سوالات ہیں۔ اس میں سے طالب علم کو کوئی پانچ سوالوں کے جواب دینے ہیں۔ ہر سوال کا جواب تقریباً دو سو (200) لفظوں پر مشتمل ہے۔
ہر سوال کے لیے 6 نمبرات مختص ہیں۔
(5 x 6 = 30 Marks)

3. حصہ سوم میں پانچ سوالات ہیں۔ اس میں سے طالب علم کو کوئی تین سوالوں کے جواب دینے ہیں۔ ہر سوال کا جواب تقریباً پانچ سو (500) لفظوں پر مشتمل ہے۔ ہر سوال کے لیے 10 نمبرات مختص ہیں۔
(3 x 10 = 30 Marks)

حصہ اول

سوال : 1

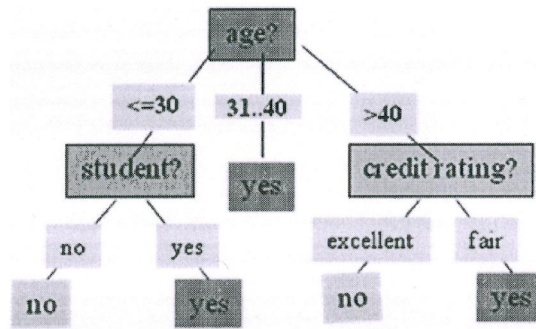
- (i) Machine Learning کی وضاحت کریں۔
- (ii) Data Mining اور Machine Learning کے مابین فرق کا ذکر کریں۔
- (iii) Machine Learning میں Overlifting کیا ہے؟
- (iv) Inductive Machine Learning کیا ہے؟
- (v) Machine Learning کی پانچ مشہور Algorithm کیا ہیں؟
- (vi) Machine Learning میں Hypotheses یا Model کو Build کرنے کے لیے تین مراحل کیا ہیں؟
- (vii) Genetic Programming کیا ہے؟
- (viii) Categorical Data کے Descriptive Analysis کے لیے دو میٹریکس Matrix لکھیں۔
- (ix) Training Set اور Test Set کیا ہیں؟
- (x) Machine Learning کے لیے مختلف طریقوں کی فہرست لکھیں۔

حصہ دوم

2. مناسب مثال کے ساتھ Bagging اور Boosting میں فرق لکھیں۔

3. Support اور Confidence میں فرق لکھیں۔

- .4 Classification اور Prediction کے حوالے سے مختلف Issues کیا ہیں؟ Back-propagation کے طریقہ کار کو استعمال کر کے Prediction کی وضاحت کریں۔ Classification اور Prediction کے کچھ Real-Life استعمال کے Applications لکھیں۔
- .5 Intelligent Model کے Performance کو کیسے Measure کیا جاتا ہے؟ Confusion Matrix کی وضاحت کریں۔
- .6 Support Vector machine (SVM) کی وضاحت کریں۔ ANN سے زیادہ SVM کے فوائد کیا ہیں؟ ANN, SVM میں Overfitting کے مسئلے پر کیسے قابو پاسکتے ہیں؟
- .7 Clustering کیا ہے؟ Clustering میں استعمال ہونے والے Algorithm پر مختصر نوٹ لکھیں۔
- .8 مندرجہ ذیل Tree-Structure پر Rule-Based Classification کو Apply کریں اور System خریدنے کے لیے Rules لکھیں۔ Missing Values کو Assume کریں۔



- .9 Hypothesis Testing کیا ہے؟ Null Hypothesis اور Alternate Hypothesis کے مابین فرق کریں۔

حصہ سوم

- .10 Classification Algorithm کو Assess اور Compare کرنے کے لیے کسی بھی دو طریقوں کو Discuss کریں۔
- Re-Sampling اور Cross-validation کے درمیان فرق کریں۔

- .11 مندرجہ ذیل Database کے Transaction کے Set کے لیے Association Rules کو ڈیزائن کریں۔

(Customer_ID:item_a1,item_a2,...).

- a: 1,3,5
b: 1, 8, 14, 17, 12
c: 4, 6, 8, 12, 9, 104.
d: 2, 1, 8.

Assuming MinSupport = 35% تمام Frequent Item Set کو تلاش کریں اور مندرجہ ذیل کو Compute کریں۔

- (i) support{8,12} (ii) support {1,5} (iii) support {1}

- .12 Ada boost Algorithm کے Steps لکھیں۔ مثال کے ساتھ ان کو بیان کریں۔ Voting Function کو Define کریں اور Bagging Algorithm کے Steps لکھیں۔

.13 مناسب مثال کے ساتھ Generative اور Discriminative Training کے مابین فرق کریں۔

- .14 Instance-based Learning کیا ہے؟ K-nearest Neighbour Algorithm کی وضاحت کریں۔