

Subject Name (Gujarati)

4341603 -- Summer 2023

Semester 1 Study Material

Detailed Solutions and Explanations

પ્રશ્ન 1(અ) [3 માકર્સ]

હુમન લર્નિંગને વ્યાખ્યાયિત કરો. હુમન લર્નિંગના પ્રકારોની યાદી બનાવો.

જવાબ

હુમન લર્નિંગ એ પ્રક્રિયા છે જેના દ્વારા માનવીઓ અનુભવ, અભ્યાસ અથવા સૂચનાઓ દ્વારા નવા જ્ઞાન, કૌશલ્યો, વર્તન મેળવે છે અથવા હાલનાઓમાં ફેરફાર કરે છે.

હુમન લર્નિંગના પ્રકારો:

પ્રકાર	વર્ણન
સુપરવાઇઝ લર્નિંગ	શિક્ષક/માર્ગદર્શકની મદદથી શીખવું
અનસુપરવાઇઝ લર્નિંગ	બાધ્ય માર્ગદર્શન વિના સ્વ-નિર્દેશિત શીખવું
રિનફોર્સમેન્ટ લર્નિંગ	ફીડબેક સાથે પ્રયાસ અને ભૂલ દ્વારા શીખવું

મેમરી ટ્રીક

“SUR - Supervised, Unsupervised, Reinforcement”

પ્રશ્ન 1(બ) [4 માકર્સ]

કવાલિટેટિવ ડેટા અને કવોનિટેટિવ ડેટા વચ્ચે તફાવત આપો.

જવાબ

ટેબલ: કવાલિટેટિવ vs કવોનિટેટિવ ડેટા

લક્ષણ	કવાલિટેટિવ ડેટા	કવોનિટેટિવ ડેટા
પ્રકૃતિ	વર્ણનાત્મક, કેટેગોરિકલ	સંઘાતાત્મક, માપી શકાય તેવું
વિશ્લેષણ	વ્યક્તિગત અર્થધારણ	આંકડાકીય વિશ્લેષણ
ઉદાહરણો	રંગો, નામો, લિંગ	ઉચ્ચાઈ, વજન, ઉંમર
પ્રતિનિધિત્વ	શબ્દો, કેટેગરીઓ	સંઘાતાત્મક, ગ્રાફ્સ

મેમરી ટ્રીક

“QUAN-Numbers, QUAL-Words”

પ્રશ્ન 1(ક) [7 માકર્સ]

મશીન લર્નિંગના વિવિધ પ્રકારોની સરખામણી કરો.

જવાબ

ટેબલ: મશીન લર્નિંગના પ્રકારોની સરખામણી

પ્રકાર	ટ્રેનિંગ ડેટા	દયેય	ઉદાહરણો
સુપરવાઇઝ	લેબલ ડેટા	પરિણામોની આગાહી	કલાસિફિકેશન, રિગ્રેશન

અનસુપરવાઈજડ
રિઝન્ફોર્મેન્ટ

અનલેબલ ડેટા
રિવર્ડ/પેનલ્ટી

પેર્ટન શોધવા
રિવર્ડ મેક્સિમાઇઝ
કરવા

કલસ્ટરિંગ, એસોસિએશન
ગેમિંગ, રોબોટિક્સ

મુખ્ય તફાવતો:

- સુપરવાઈજડ: ટ્રેનિંગ માટે ઇનપુટ-આઉટપુટ જોડીનો ઉપયોગ કરે છે
- અનસુપરવાઈજડ: ડેટામાં છુપાયેલા પેર્ટન શોધે છે
- રિઝન્ફોર્મેન્ટ: પર્યાવરણ સાથે કિયાપ્રતિક્ષિયા દ્વારા શીખે છે

મેમરી ટ્રીક

“SUR-LAP: Supervised-Labeled, Unsupervised-Reveal, Reinforcement-Action”

પ્રશ્ન 1(ક OR) [7 માકર્સ]

મશીન લર્નિંગ વ્યાખ્યાયિત કરો. મશીન લર્નિંગની કોઈપણ ચાર એપ્લિકેશનને ટૂકમાં સમજાવો.

જવાબ

મશીન લર્નિંગ આર્ટિફિશિયલ ઇન્ટેલિજન્સનો ઉપવિભાગ છે જે કમ્પ્યુટરોને સ્પષ્ટ પ્રોગ્રામિંગ વિના ડેટામાંથી શીખવા અને નિર્ણયો લેવા સક્ષમ બનાવે છે.

ચાર એપ્લિકેશનો:

એપ્લિકેશન	વર્ણન
ઈમેઇલ સ્પામ ડિટેક્શન	ઈમેઇલને સ્પામ અથવા વૈધ તરીકે વર્ગીકૃત કરે છે
ઇમેજ રેકૉર્ડિંગ	ફોટોમાં ઓળખાને આળખાને આપી છે
રેકમેન્ડશન સિસ્ટમ	યુઝરને પ્રોડક્ટ્સ/કાન્ટેન સૂચવે છે
મેડિકલ ડાયાગ્નોસ્ટિસ	રોગોની શોધમાં ડોક્ટરોની મદદ કરે છે

મેમરી ટ્રીક

“SIRM - Spam, Image, Recommendation, Medical”

પ્રશ્ન 2(અ) [3 માકર્સ]

નીચેના ઉદાહરણોનો યોગ્ય ડેટા પ્રકાર જણાવો.

જવાબ

ડેટા પ્રકાર વર્ગીકરણ:

ઉદાહરણ	ડેટા પ્રકાર
વિદ્યાર્થીઓની રાષ્ટ્રીયતા	ક્રેટેગોરિકલ (નોમિનલ)
વિદ્યાર્થીઓની શિક્ષણ સ્થિતિ	ક્રેટેગોરિકલ (ઓર્ડિનલ)
વિદ્યાર્થીઓની ઊંચાઈ	ન્યુમેરિકલ (કન્ટિન્યુઅસ)

મેમરી ટ્રીક

“NCN - Nominal, Categorical, Numerical”

પ્રશ્ન 2(બ) [4 માકર્સ]

ડેટા પ્રી-પ્રોસેસિંગ ટૂકમાં સમજાવો.

જવાબ

ડેટા પ્રી-પ્રોસેસિંગ એ મશીન લર્નિંગ અભ્યોરિધમ માટે કાચા ડેટાને તૈયાર કરવાની તકનીક છે.
મુખ્ય સ્ટેપ્સ:

સ્ટેપ	હેતુ
ડેટા કલીનિંગ	ભૂલો અને અસંગતતાઓ દૂર કરવી
ડેટા ઇન્ટીગ્રેશન	બહુવિધ સ્ત્રોતોમાંથી ડેટાને જોડવો
ડેટા ટ્રાન્સફોર્મેશન	ડેટાને યોગ્ય ફોર્મટમાં બદલવો
ડેટા રિડક્ષન	માહિતી જાળવીને ડેટાનું કદ ઘટાડવું

મેમરી ટ્રીક

"CITR - Clean, Integrate, Transform, Reduce"

પ્રશ્ન 2(ક) [7 માક્સ]

K-ફોલ્ડ કોસ વેલિડેશન વિગતવાર સમજાવો.

જવાબ

K-ફોલ્ડ કોસ વેલિડેશન એ ડેટાને K સમાન ભાગોમાં વિભાજિત કરીને મોડેલ પરફોર્મન્સ મૂલ્યાંકનની તકનીક છે.
પ્રક્રિયા:

Mermaid Diagram (Code)

```
{Shaded}
{Highlighting} []
graph LR
    A["A [ ] { -{-}{} } B [K ]"]
    B["B { -{-}{} } C [K{-}1 ]"]
    C["C { -{-}{} } D [1 ]"]
    D["D { -{-}{} } E [K ]"]
    E["E { -{-}{} } F [ ]"]
{Highlighting}
{Shaded}
```

સ્ટેપ્સ:

- વિભાજન: ડેટાસેટને K સમાન ભાગોમાં વહેંચો
- ટ્રેનિંગ: K-1 ફોલ્ડનો ઉપયોગ ટ્રેનિંગ માટે કરો
- ટેસ્ટ: બાકીના ફોલ્ડનો ઉપયોગ વેલિડેશન માટે કરો
- પુનરાવર્તન: K વખત પ્રક્રિયા કરો
- સરેરાશ: સરેરાશ પરફોર્મન્સ કાઢો

ફાયદા:

- ઓવરફિટિંગ ઘટાડે છે
- મર્યાદિત ડેટાનો બહેતર ઉપયોગ
- વધુ વિશ્વસનીય પરફોર્મન્સ અંદાજ

મેમરી ટ્રીક

"DTRA - Divide, Train, Repeat, Average"

પ્રશ્ન 2(અ OR) [3 માક્સ]

નીચેના શબ્દો વ્યાખ્યાપિત કરો: i) Mean, ii) Outliers, iii) Interquartile range

જવાબ

આંકડાકીય શબ્દો:

શબ્દ	વ્યાખ્યા
Mean	ડેટાસેટમાં બધી વેલ્યુઝની સરેરાશ
Outliers	અન્ય ડેટા પોઇન્ટ્સથી નોંધપાત્ર રીતે અલગ ડેટા પોઇન્ટ્સ
Interquartile Range	75મા અને 25મા પર્સેન્ટાઇલ વચ્ચેનો તફાવત

મેમરી ટ્રીક

"MOI - Mean, Outliers, Interquartile"

પ્રશ્ન 2(બ OR) [4 માકર્સ]

કન્ફ્યુશન મેટ્રિક્સની રૂચના સમજાવો.

જવાબ

કન્ફ્યુશન મેટ્રિક્સ સ્ક્રુચર:

આગાહી		
વાસ્તવિક	પોઝિટિવ	નેગાટિવ
પોઝિટિવ	True Positive (TP)	False Negative (FN)
નેગાટિવ	False Positive (FP)	True Negative (TN)

ઘટકો:

- TP: સાચી રીતે આગાહી કરેલા પોઝિટિવ કેસો
- TN: સાચી રીતે આગાહી કરેલા નેગાટિવ કેસો
- FP: ખોટી રીતે પોઝિટિવ તરીકે આગાહી કરેલા
- FN: ખોટી રીતે નેગાટિવ તરીકે આગાહી કરેલા

મેમરી ટ્રીક

"TTFF - True True, False False"

પ્રશ્ન 2(ક OR) [7 માકર્સ]

ફીચર સબસેટની પસંદગી પર ટૂંકી નોંધ લખો.

જવાબ

ફીચર સબસેટ સિલેક્શન એ મૂળ ફીચર સેટમાંથી સંબંધિત ફીચર્સ પસંદ કરવાની પ્રક્રિયા છે.

મેથ્ડ્સ:

મેથ્ડ	વર્ણન
ફિલ્ટર મેથ્ડ્સ	ફીચર્સ રેન્ક કરવા આંકડાકીય માપદંડોનો ઉપયોગ
રેપર મેથ્ડ્સ	ફીચર સબસેટ્સ મૂલ્યાંકન માટે ML અલ્ગોરિધમનો ઉપયોગ
એમ્બેડેડ મેથ્ડ્સ	મોડેલ ટ્રેનિંગ દરમિયાન ફીચર સિલેક્શન

ફાયદા:

- ઘટાડેલી જટિલતા: ઓછા ફીચર્સ, સરળ મોડેલ્સ
- સુધારેલ પરકોર્પન્સ: નોઇજ અને અપ્રસ્તુત ફીચર્સ દૂર કરે છે
- જડપી ટ્રેનિંગ: ઓછો કમ્પ્યુટેશનલ ઓવરહેડ

લોકપ્રિય તકનીકો:

- Chi-square ટેસ્ટ
- Recursive Feature Elimination
- LASSO રેઝ્યુલરાઇઝેશન

મેમરી ટ્રીક

"FWE - Filter, Wrapper, Embedded"

પ્રશ્ન 3(અ) [3 માંકર્સ]

પ્રેડિક્ટિવ મોડેલ અને ડીરિક્ટિવ મોડેલ વચ્ચેનો તફાવત આપો.

જવાબ

મોડેલ પ્રકાર સરખામણી:

લક્ષણ	પ્રેડિક્ટિવ મોડેલ	ડીરિક્ટિવ મોડેલ
હેતુ	ભાવિ પરિણામોની આગાહી	વર્તમાન પેટર્ન સમજવા
આઉટપુટ	આગાહીઓ/વર્ગીકરણ	અંતર્દૃષ્ટિ/સારાંશ
ઉદાહરણો	રિચ્રેશન, કલાસિફિકેશન	કલસ્ટરિંગ, એસોસિએશન રૂલ્સ

મેમરી ટ્રીક

"PF-DC: Predictive-Future, Descriptive-Current"

પ્રશ્ન 3(બ) [4 માંકર્સ]

કલાસિફિકેશન અને રિચ્રેશન વચ્ચેના તફાવતની ચર્ચા કરો.

જવાબ

કલાસિફિકેશન vs રિચ્રેશન:

પાસું	કલાસિફિકેશન	રિચ્રેશન
આઉટપુટ	ડિસ્ક્રીટ કેટેગરીઓ	કન્ટિન્યુઅસ વેલ્યુઝ
ધ્યેય	કલાસ લેબલ્સની આગાહી	ન્યૂમેરિકલ વેલ્યુઝની આગાહી
ઉદાહરણો	સ્પામ ડિટેક્શન, ઇમેજ રેકૉર્ડિંગ	કિમત આગાહી, તાપમાન
મૂલ્યાંકન	Accuracy, precision, recall	MSE, RMSE, R-squared

મેમરી ટ્રીક

"CCNM - Classification-Categories, Regression-Numbers"

પ્રશ્ન 3(ક) [7 માંકર્સ]

કલાસિફિકેશનને વ્યાખ્યાપિત કરો. કલાસિફિકેશન લર્નિંગના સ્ટેપને વિગતોમાં સમજાવો.

જવાબ

કલાસિફિકેશન એ સુપરવાઇઝ લર્નિંગ તકનીક છે જે ઇનપુટ ડેટા માટે ડિસ્ક્રીટ કલાસ લેબલ્સની આગાહી કરે છે.
કલાસિફિકેશન લર્નિંગ સ્ટેપ્સ:

Mermaid Diagram (Code)

```
{Shaded}
{Highlighting} []
graph LR
    A[ ] --- B[ ]
    B --- C[ ]
    C --- D[ { } ]
    D --- E[ ]
    E --- F[ ]
    F --- G[ ]
{Highlighting}
{Shaded}
```

વિગતવાર સ્ટેપ્સ:

- ડેટા કલેક્શન: લેબલ ટ્રેનિંગ ડેટા એકત્ર કરવો
- પ્રીપોરેસિંગ: ડેટાને સાફ કરવો અને તૈયાર કરવો
- ફીચર સિલેક્શન: સંબંધિત લક્ષણો પસંદ કરવા
- ડેટા સ્પલટ: ટ્રેનિંગ અને ટેસ્ટિંગ સેટમાં વિભાજન
- ટ્રેનિંગ: ટ્રેનિંગ ડેટાનો ઉપયોગ કરીને મોડેલ બનાવવું
- મૂલ્યાંકનાં: મોડેલ પરફોર્મન્સ ચકાસવી
- ડિલોયમેન્ટ: આગાહીઓ માટે મોડેલનો ઉપયોગ

મેમરી ટ્રીક

“DCFSTED - Data, Clean, Features, Split, Train, Evaluate, Deploy”

પ્રશ્ન 3(અ OR) [3 માકર્સ]

બેગિંગ અને બૂસ્ટિંગ વચ્ચેનો તફાવત આપો.

જવાબ

બેગિંગ વિનિયોગ વિનિયોગ:

લક્ષણ	બેગિંગ	બૂસ્ટિંગ
સેમ્પલિંગ	બૂટસ્ટ્રેપ સેમ્પલિંગ	કમાનુગત વેઇટેડ સેમ્પલિંગ
ટ્રેનિંગ	પેરેલલ ટ્રેનિંગ	કમાનુગત ટ્રેનિંગ
ફિક્સ	વેરિયન્સ ઘટાડવું	બાયસ ઘટાડવું

મેમરી ટ્રીક

“BPV-BSB: Bagging-Parallel-Variance, Boosting-Sequential-Bias”

પ્રશ્ન 3(બ OR) [4 માકર્સ]

લોજિસ્ટિક રિગ્રેશનના વિવિધ પ્રકારો સંક્ષિપ્તમાં સમજાવો.

જવાબ

લોજિસ્ટિક રિગ્રેશનના પ્રકારો:

પ્રકાર	કલાસો	ઉપયોગ
બાઇનરી	2 કલાસો	હ/ના, પાસ/ફેઇલ

મલ્ટિનોમિયલ 3+ કલાસો (અવ્યવસ્થિત)
ઓર્ડિનલ 3+ કલાસો (ક્રમાંકિત) રંગ વર્ગીકરણ
રેટિંગ સ્કેલ

મેમરી ટ્રીક

“BMO - Binary, Multinomial, Ordinal”

પ્રશ્ન 3(ક) OR) [7 માકર્સ]

k-NN અલ૗ોરિધમ લખો અને તેના ઉપયોગ બતાવો.

મેમરી ટ્રીક

“CDSA - Choose, Distance, Select, Assign”

પ્રશ્ન 4(અ) [3 માકર્સ]

સપોર્ટ વેક્ટર મશીનની એપ્લિકેશનોની ચાદી બનાવો.

જવાબ

SVM એપ્લિકેશનો:

એપ્લિકેશન	ડોમેન
ટેક્સ્ટ કલાસિફિકેશન	ડોક્યુમેન્ટ કેટેગોરાઇઝેશન
ઇમેજ રેકૉર્ડિંગ	ફેસ ડિટેક્શન
બાયોઇન્ફોર્મેટિક્સ	જીન કલાસિફિકેશન

મેમરી ટ્રીક

“TIB - Text, Image, Bio”

પ્રશ્ન 4(બ) [4 માકર્સ]

k-means અલ૗ોરિધમ માટે સ્યુડો કોડ બનાવો.

જવાબ

K-means સ્યુડો કોડ:

```
BEGIN K-means
1. k
2. REPEAT
    a.
    b.
    c. -
3. UNTIL
4. RETURN
END
```

મેમરી ટ્રીક

“IAUC - Initialize, Assign, Update, Check”

પ્રશ્ન 4(ક) [7 માકર્સ]

અનસુપરવાઇઝ લર્નિંગની એપ્લિકેશનો લખો અને સમજાવો.

જવાબ

અનસુપરવાઇઝ લર્નિંગ લેબલ ઉદાહરણો વિના ડેટામાં છુપાયેલા પેટર્ન શોધે છે.

મુખ્ય એપ્લિકેશનો:

એપ્લિકેશન	વર્ણન	ઉદાહરણ
કસ્ટમર સેગમેન્ટેશન	વર્તન પ્રમાણો ગ્રાહકોનું ગુપ્તિંગ	માર્કેટ રિસર્ચ
એનોમેલી ડિટેક્શન	અસામાન્ય પેટર્ન ઓળખવા	ફોડ ડિટેક્શન
ડેટા કમ્પોશન	ડાયમેન્શનાલિટી ઘટાડવી	ઇમેજ કમ્પોશન
એસોસિએશન રૂલ્સ	આઇટમ સંબંધો શોધવા	માર્કેટ બાસકેટ વિશ્લેષણ

કલસ્ટરિંગ એપ્લિકેશનો:

- માર્કેટ રિસર્ચ: કસ્ટમર ગુપ્તિંગ
- સોશિયલ નેટવર્ક વિશ્લેષણ: કમ્યુનિટી ડિટેક્શન
- જીન સીકવેન્સિંગ: બાયોલોજિકલ કલાસિફિકેશન

ડાયમેન્શનાલિટી રિડક્શન:

- વિશ્યુઅલાઇઝેશન: હાઇ-ડાયમેન્શનલ ડેટા પ્લોટિંગ
- ફીચર એક્સ્પ્રેક્શન: નોઇડ રિડક્શન

મેમરી ટ્રીક

"CADA - Customer, Anomaly, Data, Association"

પ્રશ્ન 4(અ OR) [3 માકર્સ]

રિશેનની એપ્લિકેશનોની યાદી બનાવો.

જવાબ

રિશેન એપ્લિકેશનો:

એપ્લિકેશન	હેતુ
સ્ટોક પ્રાઇસ પ્રેડિક્શન	ફાઇનાન્શિયલ ફોરકાસ્ટિંગ
સેલ્સ ફોરકાસ્ટિંગ	બિઝનેસ પ્લાનિંગ
મેડિકલ ડાયગ્નોસ્ટિસ	રિસ્ક એસેસમેન્ટ

મેમરી ટ્રીક

"SSM - Stock, Sales, Medical"

પ્રશ્ન 4(બ OR) [4 માકર્સ]

નીચેના શર્દી વ્યાખ્યાયિત કરો: i) Support ii) Confidence

જવાબ

એસોસિએશન રૂલ શર્દી:

શર્દી	વ્યાખ્યા	ફોર્મ્યુલા
Support	ડેટાબેઝમાં આઇટમસેટની આવર્તન	Support(A) = A / D
Confidence	રૂલની શરતી સંભાવના	Confidence(A) = Support(A) / Support(A)

ଓଡାହରଣ:

- જો 30% ટ્રાન્జેક્શનમાં બેડ અને દૂધ હોય: Support = 0.3
 - જો 80% બેડ ખરીદનારાઓ દૂધ પણ ખરીદે: Confidence = 0.8

ਮੇਮਰੀ ਟ੍ਰੀਕ

“SF-CP: Support-Frequency, Confidence-Probability”

प्रश्न 4(क) OR [7 मार्कसे]

apriori algorithm ने विगतवार समजावो.

જવાબ

Apriori અંગ્રોરિધમ એપ્રિયોરી પ્રોપર્ટીનો ઉપયોગ કરીને ટ્રાન્జેક્શનલ ડેટામાં ફીકવન્ટ આઇટમસેટ્સ શોધે છે.

અલોરિધમ સ્ટેપ્સ:

Mermaid Diagram (Code)

```
{Shaded}
{Highlighting}[]
graph LR
    A[      1{-}      ] {--{-}{-}{-}} B[      2{-}      ]
    B {--{-}{-}{-}} C[      ]}
    C {--{-}{-}{-}} D[      ]}
    D {--{-}{-}{-}} E[      k{-}      ]}
    E {--{-}{-}{-}} F\{      ?\}\}
    F {--{-}{-}{-}|} | B}
    F {--{-}{-}{-}|} | G[      ]}
{Highlighting}
{Shaded}
```

એપ્રિઅરી પ્રોપર્ટી:

- જો આઇટમસેટ ફીકવન્ટ છે, તો તેના બધા સબસેટ્સ ફીકવન્ટ છે
 - જો આઇટમસેટ ઇનફીકવન્ટ છે, તો તેના બધા સપરસેટ્સ ઇનફીકવન્ટ છે

स्टेप्सः

1. ડેટાબેઝ સ્કેન: 1-આઇટમ સપોર્ટ કાઉન્ટ કરો
 2. કેન્દ્રિકેટ્સ જનરેટ: ફીકવન્ત k-આઇટમસેટ્સમાંથી k+1 આઇટમસેટ્સ બનાવો
 3. પ્રૂણ: ઇનફીકવન્ત સબસેટ્સ સાથેના કેન્દ્રિકેટ્સ દૂર કરો
 4. સપોર્ટ કાઉન્ટ: કેન્દ્રિકેટ ફીકવન્તી માટે ડેટાબેઝ સ્કેન કરો
 5. પુનરાવર્તન: નવા ફીકવન્ત આઇટમસેટ્સ ન મળે ત્યાં સુધી

એપ્લિકેશનો:

- માર્કેટ બાસ્કેટ વિશ્વલેખાણ
 - વેબ યુઝર પેર્ટન
 - પ્રોટીન સીકવન્સ

ਮੇਮਰੀ ਟੀਕ

“SGPCR - Scan, Generate, Prune, Count, Repeat”

પ્રશ્ન 5(અ) [3 માંકર્સ]

matplotlib ના મુખ્ય ફીચર્સની ચાદી બનાવો.

ଜୟାମ୍

Matplotlib ફીચર્સ:

ફીચર	વર્ણન
માલ્ટિપલ પ્લોટ ટાઇપ્સ	લાઇન, બાર, સ્કેટર, હિસ્ટોગ્રામ
કસ્ટમાઇઝેશન	કલર્સ, સ્ટાઇલ્સ, લેબલ્સ
એક્સપોર્ટ ઓપ્શન્સ	PNG, PDF, SVG ફોર્મેટ્સ

મેમરી ટ્રીક

“MCE - Multiple, Customization, Export”

પ્રશ્ન 5(બ) [4 માકર્સ]

Numpy ના પ્રોગ્રામમાં iris ડેટાસેટ કેવી રીતે લોડ કરવો? સમજાવો.

જવાબ

NumPy માં Iris ડેટાસેટ લોડ કરવું:

```
import numpy as np
from sklearn.datasets import load_iris

# iris
iris = load_iris()
data = iris.data    #
target = iris.target #
```

સ્ટેપ્સ:

- Import: જરૂરી લાઇબ્રેરીઓ import કરો
- Load: sklearn ના load_iris() ફંક્શનનો ઉપયોગ કરો
- Extract: ફીચર્સ અને ટાર્ગેટ એરે મેળવો
- Access: .data અને .target એટ્રિબ્યુટ્સનો ઉપયોગ કરો

મેમરી ટ્રીક

“ILEA - Import, Load, Extract, Access”

પ્રશ્ન 5(ક) [7 માકર્સ]

Pandas ની વિશેષતાઓ અને એપ્લિકેશનો સમજાવો.

જવાબ

Pandas એ Python માટે શક્તિશાળી ડેટા મેનિપ્યુલેશન અને વિશ્લેષણ લાઇબ્રેરી છે.

મુખ્ય ફીચર્સ:

ફીચર	વર્ણન
DataFrame	2D લેબલ ડેટા સ્ટ્રક્ચર
Series	1D લેબલ એરે
Data I/O	વિવિધ ફાઈલ ફોર્મેટ્સ વાંચવા/લખવા
Data Cleaning	બિસિંગ વેલ્યુઝ હેન્ડલ કરવા
Grouping	ગુપ અને એગ્રીગેટ ઓપરેશન્સ

એપ્લિકેશનો:

એપ્લિકેશન	ઉપયોગ
ડેટા એનાલિસિસ	આંકડાકીય વિશ્લેષણ
ડેટા કલીનિંગ	ML માટે પ્રીપ્રોસેસિંગ
ફાઇનાન્શિયલ એનાલિસિસ	સ્ટોક માર્કેટ ડેટા
વેબ સ્ક્રેપ્ટિંગ	HTML ટેબલ્સ પાર્સ કરવા

સામાન્ય ઓપરેશન્સ:

- ડેટા વાંચવો: pd.read_csv(), pd.read_excel()
- ફિલ્ટરિંગ: df[df['column'] > value]
- ગ્રુપિંગ: df.groupby('column').mean()

મેમરી ટ્રીક

“DSDCG - DataFrame, Series, Data I/O, Cleaning, Grouping”

પ્રશ્ન 5(અ OR) [3 માક્ર્સ]

matplotlib ની એપ્લિકેશનોની યાદી બનાવો.

જવાબ

Matplotlib એપ્લિકેશનો:

એપ્લિકેશન	હેતુ
સાયન્ટિફિક વિગ્યુઅલાઇઝેશન	રિસર્ચ ડેટા પ્લોટિંગ
બિઝનેસ એનાલિટિક્સ	ડેશબોર્ડ બનાવવું
એજ્યુકેશનલ કન્ટેન્ટ	શિક્ષણ સામગ્રી

મેમરી ટ્રીક

“SBE - Scientific, Business, Educational”

પ્રશ્ન 5(બ OR) [4 માક્ર્સ]

Pandas માં CSV ફાઇલ ઇમ્પોર્ટ કરવાના સ્ટેપ્સ લખો અને સમજાવો.

જવાબ

Pandas માં CSV ઇમ્પોર્ટ કરવાના સ્ટેપ્સ:

```
import pandas as pd

# 1: pandas      import
# 2: read\csv()
df = pd.read\csv({filename.csv})

#
df = pd.read\csv({file.csv},
                 header=0,      #
                 sep={,},       #
                 index\col=0)  #
```

પ્રક્રિયા:

- Import: pandas લાઇબ્રેરી import કરો
- Read: pd.read_csv() ફંક્શનનો ઉપયોગ કરો
- Specify: ફાઇલ પાથ અને પેરામીટર્સ ઉમેરો

- **Store:** DataFrame વેરિએબલમાં અસાઇન કરો

મેમરી ટ્રીક

“IRSS - Import, Read, Specify, Store”

પ્રશ્ન 5(ક) OR) [7 માંકર્સ]

Scikit-Learn ની વિશેષતાઓ અને એપ્લિકેશનો સમજાવો.

જવાબ

Scikit-Learn એ Python માટે વ્યાપક મશીન લર્નિંગ લાઇબ્રેરી છે.

મુખ્ય ફીચર્સ:

ફીચર	વર્ણન
અલોરિધમ્સ	કલારિફિકેશન, રિગ્રેશન, કલસ્ટરિંગ
પ્રીપોરેસિંગ	ડેટા સ્કેલિંગ અને ટ્રાન્સફોર્મેશન
મોડેલ સિલેક્શન	કોસ-વેલિડેશન અને ગ્રિડ સર્ચ
મેટ્રિક્સ	પરફોર્મન્સ મૂલ્યાંકન ટૂલ્સ

એપ્લિકેશનો:

ડોમેન	ઉપયોગ
હેલ્થક્રે	રોગ આગાહી
ફાઇનાન્સ	કેર્ડિટ સ્કોરિંગ
માર્કેટિંગ	કસ્ટમર સેઝમેન્ટેશન
ટેકનોલોજી	રેકમેન્ડેશન સિસ્ટમ્સ

અલોરિધમ કેટેગેરીઓ:

- સુપરવાઈઝ: SVM, Random Forest, Linear Regression
- અનસુપરવાઈઝ: K-means, DBSCAN, PCA
- એન્સેમ્બલ: Bagging, Boosting

વર્કફ્લો:

1. ડેટા તૈયારી: પ્રીપોરેસિંગ
2. મોડેલ સિલેક્શન: અલોરિધમ પસંદ કરો
3. ટ્રેનિંગ: ડેટા પર મોડેલ ફિટ કરો
4. મૂલ્યાંકન: પરફોર્મન્સ આકારો
5. આગાહી: ફોરકાસ્ટ બનાવો

મેમરી ટ્રીક

“APME - Algorithms, Preprocessing, Metrics, Evaluation”