Name : Rashika Khatri

Roll no. : 1918594

Section : B

Subject : AI

Subject Code : TCS 421

Sign : Rashika

1) a)

Application Areas of AI

- AI in finance

→ Guide people in investing money in share market based on analysis of history. We can develop a system which predict the behaviour of market based on previous experience.

S1. Collecting Data of past 20-30 years.

→ Fraud Transaction: Here, AI can play a very important role by checking whether the transaction is fraud or not by collecting data of past years & checking in which scenario the transaction is fraud. While providing intelligence to the machine here data gathering and data cleaning is very important for fraud detection, we must have thousands of sample by which we can get an eq or hypothesis and further the values we get in those parameters, we can take decision that the transaction is fraud or not. These modals will be more efficient than human calculations.

→ Personalised Banking: In banks we do not know about the various activities or bank related tasks, so here chatbots can work as a banking assistant.

- AI in healthcare

→ One of the emerging research sector where we can apply AI based algorithms and AI based modelling.

→ Medical Image Analysis: For eg: To identify the fracture on any particular part of the body, image processing plays an important role.
There might be a possibility that based on certain past history, the AI based modal can automatically analyze the particular thing like X-rays, and generate report.

Digital consultant

→ Robotic Surgery: If we have some expert system which can guide the junior doctors

→ Robotic surgery: For this, we must need to achieve almost nearby 100% accuracy for performing that task. If a robot is able to perform similar type of operations with good accuracy & efficiency just as the expert doctor, then we can say that our machine is performing similar task as a human expert can perform.
Also the robot must be able to learn from the environment.

• Education System

→ Simplifying the administrative task: To generate the report of the students on the basis of their overall performance, and to categorize them on the basis of their ranking. To create cluster of weak students we can apply clustering algorithm.

→ Student Counselling: We can create a chatbot which can help students for counselling, by collecting data of past 10 yr pass out students. Based on their marks we can create a report of profession they have chosen after graduation.

- Agriculture :

→ ~~Fast~~ weather forecast : Machines can be developed which tells the weather report of present day or upcoming 2-3 days.

→ Intelligent machines can be developed which can apply pesticides by just visualizing that particular plant or leaf, and based on the dipcase.

→ Monitoring of agricultural land : Smart cameras or smart devices to check how much agricultural land is available, and how much is used.

- Transport

→ smart Traffic Management : For emergency vehicles we can adjust the traffic using visual methods Giving directions while a person is driving.

- Robotics

→ Erica & Shophia : can react like a human. They have expressive way of talking

3) a) Data Science is a blend of various tools, algorithms, and machine learning principles with the goal to ~~discuss~~ discover hidden patterns from the raw data.

Role of data scientist :-
Data scientists work closely with business stakeholders to understand their goals and determine how data can be used to achieve those goals. They design data modelling processes, create algorithms & predictive methods to extract the data the business needs, and help analyze the data and share insight with pers.

Path they follow is :
1. Ask the right quel to begin discovery process.
2. Acquire data
3. Process the data and clean data.
4. Interpret & store data.
5. Initial data investigation & exploratory data analysis.
6. Choose one or more models & algorithms.
7. Present find result to stakeholders.

8. Measure & improve results

Different machine learning algorithms :

i) Linear Regression :
- Straight enough condition (linearity)
- Homoscedasticity - Variance of residual i/p same for $\forall x$.
- Normality. For any fixed value of x, Y i/p normally distributed.

ii) Logistic Regression ⊖
- Independent of each other.
- No multicolinearity among independent variable.
- Linearity in logit for continous variable

iii) Native Byas :
- It assumes that the presence of a particulas feature in a class is unrelated to the presence of any other feature.

iv) KNN Algorithm :

- used for classification.
- when we want to know how its neighbour is classified.

v) K-mean Algorithm :

- When we have to minimize the sum of distance b/w the points and their respective clusters centroid.

vi) SVM Algorithm :

- Used for margin maximization

vii) PCA Algorithm :

- Variables which are strongly correlated.

5) a)

Logistic	Linear
• It is used to predid the categorical dependent variable using a given set of independent variable.	It is used to predict the continous dependent variable using a given set of independent variables
• It is used for solving classification problem.	It is used for solving regression problem.
• In this, we predict the values of categorical variable	In this, we predict the ~~variab~~ value of continous variable
• Maximum likelihood estimato is used for accuracy	Least square estimation method is used for accuracy.
• Output must ~~of~~ be categorical value such as 0/1, Yes/No, etc.	Output must be continous value such as price, age,

There may not be collinearity	There may be collinearity.

In machine learning

Logistic - Supervised learning
Linear - Supervised learning

Justification : Because it uses true labels for training. & value is based on independent variables.

Eg :

- Logistic Regression
  - Credit Card Fraud : When a credit card transaction happens, the bank makes a note of several factors. For instance, the date, amount, place, types, etc.

Based on these, they develop a logistic regression model of whether or not the transaction is fraud.

For eg: If the amount is too high & the bank knows the concerned person never purchase that high, it may be label as fraud.

linear regression example:

Year	Sales (million euro)	Advertising (million euro)
1	651	21
2	762	26
3	856	30
4	1063	34
5	1190	43
6	1298	48
7	1421	52
8	1440	57
9	1518	58

The table shows clothing company data. Each row show sales for a year & an amount on advertising.

In this case, outcome of interest is sales- it is what we want to predict-

Advertising as predictor variable.

linear regression estimates:

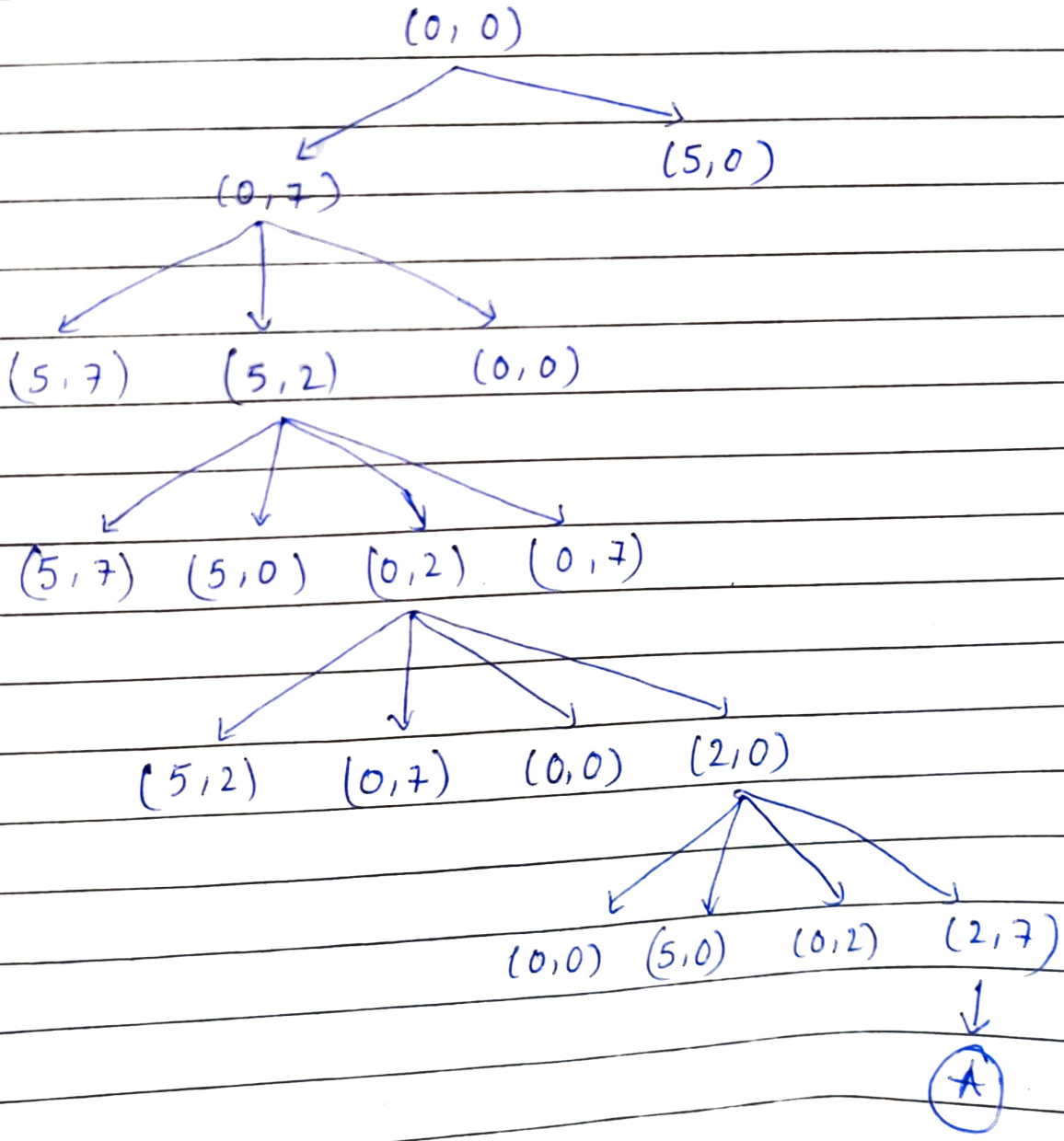
Sales = 168 + 23 Advertising,

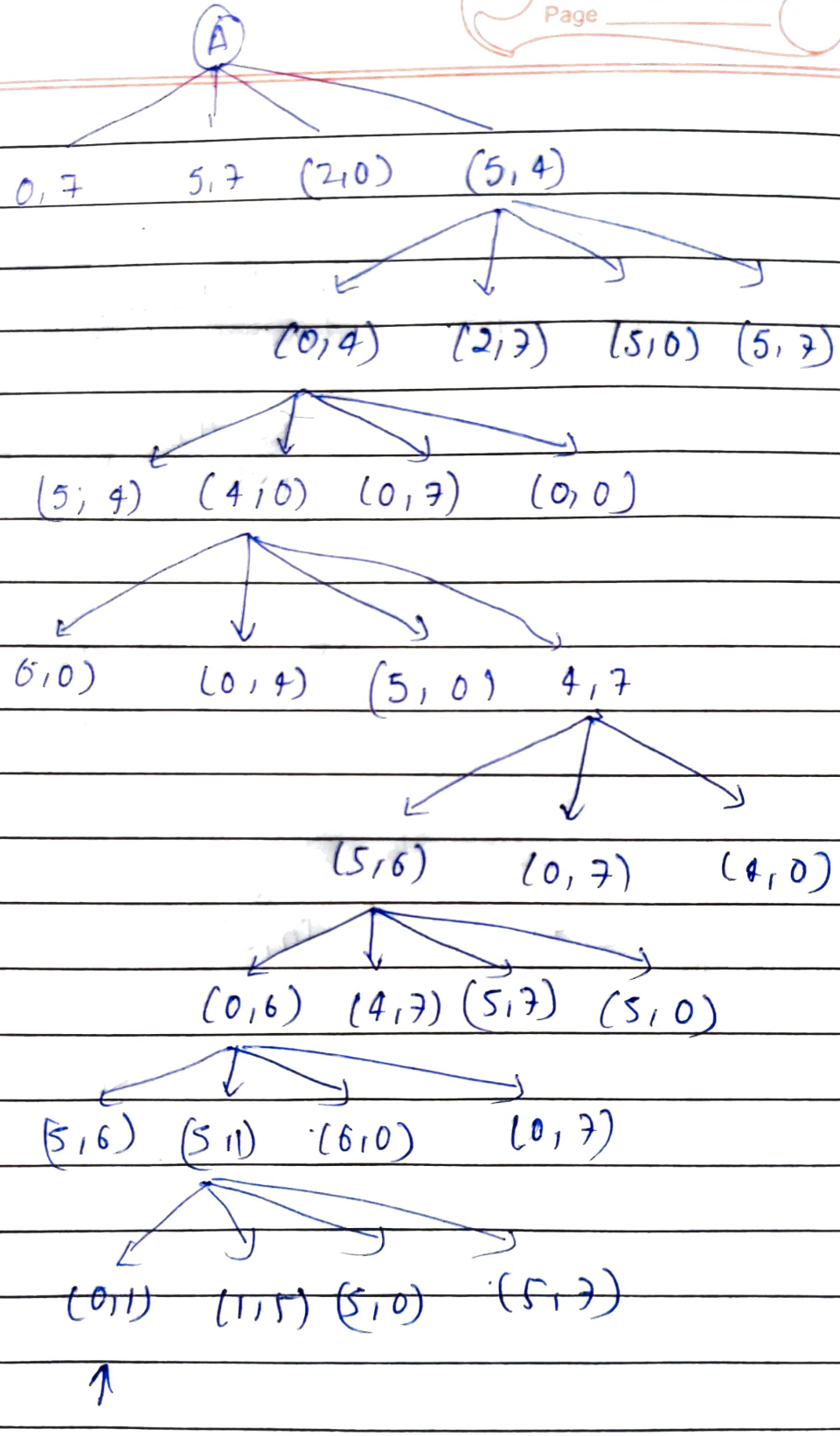
ie, if a advertising expenditure is increased by 1 million euro, then sales will be expected to increased by 23 million euros

2b) Initial state : (0, 0)

Final state : (1, 0) or (0, 1)

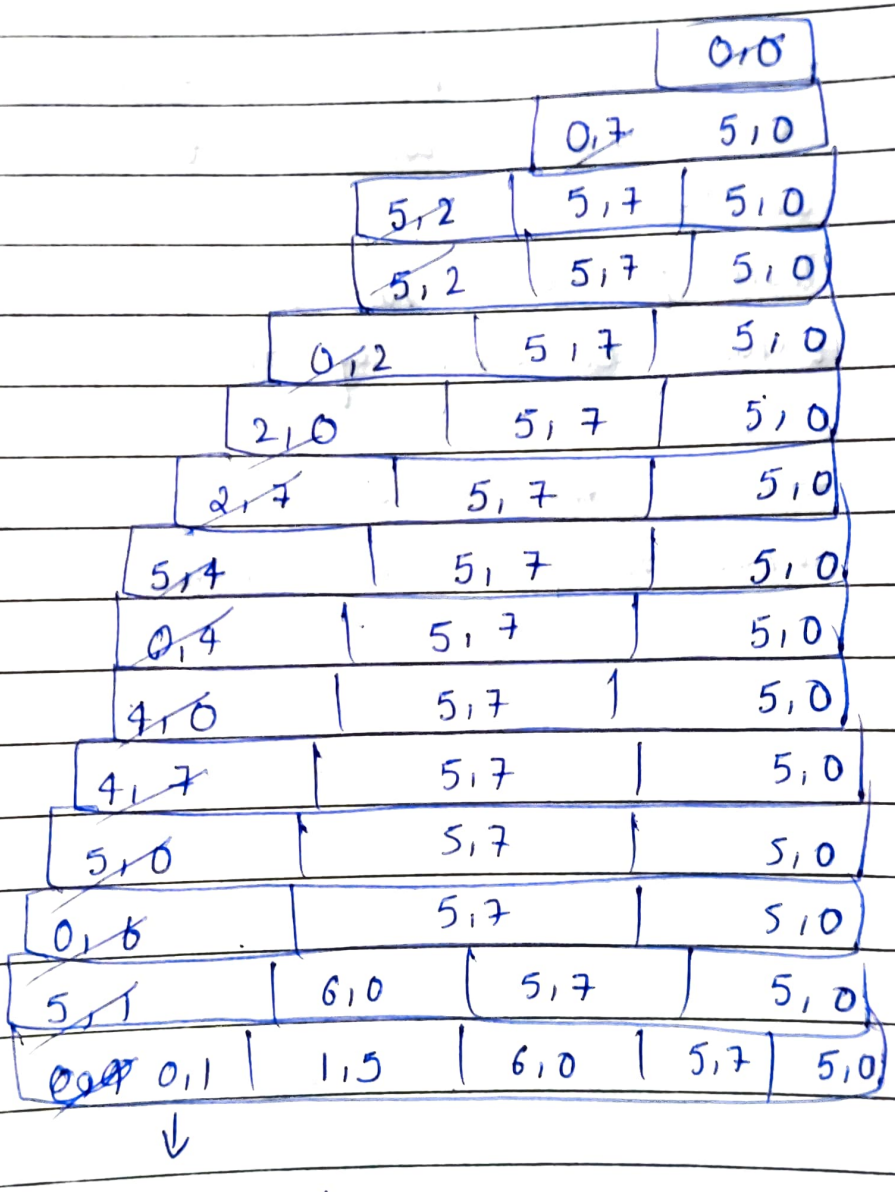
State Space Representation:




A
0,7
5,7
(2,0)
(5,4)
(0,4)
(2,7)
(5,0)
(5,7)
(5,4)
(4,0)
(0,7)
(0,0)
6,0)
(0,4)
(5,0)
4,7
(5,6)
(0,7)
(4,0)
(0,6)
(4,7)
(5,7)
(5,0)
(5,6)
(5,1)
(6,0)
(0,7)
(0,1)
(1,5)
(5,0)
(5,7)

Goal Node.

DFS Approach



↓

Goal node

Algorithm DFS:

- Insert node value in the stack
- then check whether it is a goal node or not.
- Then if it is not a goal node then add other children of the node in the stack.
- And if it is the goal node the stop and we get the answer

4) a) I) Narrative rate = 10

i) $x = 4$, using Poisson Distribution

$$P(x=4) = \frac{\lambda^n e^{-\lambda}}{x!}$$

$$= \frac{(10)^4 e^{-10}}{4!}$$

$$P(x=4) = 0.01891$$

ii) $x = 12$, $\lambda = ?$

$$\frac{\lambda_2}{\lambda_1} = \frac{5}{1}$$

$$\boxed{\lambda_2 = 50}$$

$$P(x=12) = \frac{\lambda^n e^{-\lambda}}{x!}$$

$$P(x=12) = \frac{(50)^{12} e^{-50}}{12!}$$

II) Using Binomial Distribution,

$n = 400, \quad P = 0.01, \quad q = 0.99$

$P(x \geq 5) = 1 - P(x < 5)$

$P(x=0) = {}^{400}C_0 \, (0.01)^0 \, (0.99)^{400}$

$= 0.01795$

$P(x=1) = {}^{400}C_1 \, (0.01)^1 \, (0.99)^{399}$

$= 0.07252$

$P(x=2) = {}^{400}C_2 \, (0.99)^{398}$

$= 200 \times 399 \times (0.1)^2 \times (0.99)^{398}$

$= 0.14615$

$P(x=3) = {}^{400}C_3 \, (0.01)^3 \, (0.99)^{397}$

$= \frac{\overset{20}{\cancel{400}} \times 399 \times \overset{133}{\cancel{399}}}{\cancel{3} \times \cancel{2}}$

$= 0.19585$

$P(X=4) = {}^{400}C_4 (0.01)^4 (0.99)^{396}$

$= 0.19635$

$P(X \geq 5) = 1 - P(X < 5)$

$1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3) + P$

$= 1 - [0.01795 + 0.07252 + 0.14615 + 0.19585 + 0.19635]$

P (atleast 5 are defective) $= 0.37118$