	military is the same of the sa
0	Bank (BDA):
	and a series of the series of
0.	What is Data Streaming and explain with example.
0 1	
7	Vata streaming in the context of BDA refers to the
	Pata streaming in the context of BDA refers to the continuous real-time processing and analysis of large volumes of Jata as they are generated.
(40 P)	volumes of data as they are generated.
	Unlike traditional batch processing where data is collected
	Stores and ohen processed at a later time, data
	Streaming processes data on-the-thy as it arrives,
,	Unlike traditional batch processing, where data is collected stored and then processed at a later time, data streaming processes data on-the-fly as it arrives, enabling immediate insights and responses.
1	- xample:
3 5 7 6 .	Example: Data Streaming in Sontiment Analysis on Social Media. Where posts and comments about a brand are monitored
	where posts and comments about a brand are monitored
- 1	in a la film of
1	Ising technologies like Apache Kafka for data ingestion and Apache Flink or Spark Streaming for data processing, companies can analyse sentiments shrough natural language processing to measure public opinion.
	Anacha Fink or know Streaming for date
	macaccina com annes can an luca salatiments domanan
	Lucil language proceeding to maggine hubble contran
	hatosal language poolessing to measure poole opinion.
40.0	The state of the s
	The state of the s
-	The state of the s

÷	DATE / /
Q.	Explain Content-based Recommendations
	16921 Sugarini
7	Sontent-based recommendation systems are a type of recommendation system that suggest items to
	usors based on the features of the items and
offs.	a profile of their user's preferences.
16/ V	c sistinco fue dissolute vide con and unitsis o
	Content-based recommendation are widely used in various
	domains, such as recommending movies, books, news
131180	articles and more tailored to the individual's
Divisor	explicit préférences.
	Advantages:
	· Lexionalisation:
binso.	Recommendations are tailored to the individuals
	specific preferences, based on the content they have engaged with
	navo engageo wioh, sold sold sold sold sold sold sold sold
	· No Pold Start for item:
	Since recommendations are based on item
dings.	features, neuritems can be recommended as soon as
· idia:	Shey are available.
	M Challenges:
	· Usor- Profile Initialization -> New User = 10 fistory
	and sury as a beginner are needs to
	Ohallenges: - Usar-Profile Initialization -> New User = No History Short's why as a beginner was needs to go Shrough direct queries about preference
	· Over-specialization -> User ko same-same content
	repeat he sakta jisti wajah se user haqidirmity
	· Over-specialization -> User ko same-same content repeat ho sakta jisli wajah se user hagi direnity discover rahi kar poyega

.bs

	PAGE No. DATE //
O.	Explain Relational Operation using Map Reduce.
→	Mapleduce is a programming model designed to process large volumes of data in a distributed computing environment
	It consists of two main phases: The Map Phase & The Raduca Phase
	Despite being primarily designed for simple aggregation dasks, such as counting occurrence of words in a large set of documents, Map Reduce can also adopted to perform more country and applications
	to perform more complex operations such as those found in relational databases
	Relational Operations in MapRoduce:
Hagragation)	I Filtering (Selection): Imagine it you have a long list of people with details like age & you only want to keep the details of people over 18. In the map step, you go through each record and lick out only those that say "age > 18".
hroop J	The reduce step, you just collect all these iltered records together.
١	Think of this as choosing only the columns you are storested infrom a table. If you have a table with columns like Name, Age &
G	by Not you prly care about Names, in oh Map 5 top
da da	The Reduce step might not even be needed unless you want to something extra

	PRICE NO.
Q.	How Bloom filter is useful for his data analytics
Q.	How Bloom filter is useful for big data analytics. Explain with one example.
1KN	the control of the cost forth was not an of sould like these
0 2->	A Bloom filter is a space-efficient probabilistic
)	A Bloom filter is a space-efficient probabilistic data structure that is used to test whether an element is a member of a set.
	element is a member of a set.
F. 050	It will also esende ainer and the string the
3 3	It's particularly useful in situations where saying saving memory is crucial and a small probability of error can be tolerated.
	saving memory is crucial and a small probability
nch	cotrerror can be tolorated.
120	Energy to survey to the form the first
Idepa	For Big data analytics; where datasets can be extraordinarily large and checking each element against a dataset or database can be time-
1101	extraordinarily large and checking each element
	against a dataset or jotabase can be time-
	offer a fast & officient way to redoce the
	hered for expensive lookups or checks.
	There is the constant
6-1-1	Advantages: Space Efficient
To	Timo Efficient de la companya della
	The same of the sa
345 8	Example: Web Krawling
	Tollo righ sort sont short son son sold
67 %	Let's say you are building a web crawler to de visit
	different Inh Dage
	You want to avoid revisiting pages you have already
	SULV 50 200 0 010 000 0000 0000 0000
UC W	A Bloom filter can quickly tell you it you have
	massive list of every we page you've ever seen
	massive list of every we have you've know seen
a.	and ah di said all the said and the
	Livery was at the control of the second

त्र क्षेत्रं स क्ष

2/1

Fig. West States

	PAGE No. DATE //
Q.	Explore the clusting application
-7	Olustoring is a fundamental techique in data analysis and machine learning, used to group similar data-points together based on their characteristics or features.
	1 lustomer Segmentation: . E-lommerce -> lystering helps businesses understand
	Customers with similar purchasing patterns. Unko Ads dikhate hai same!
	E-lammerce -> Customer behaviour by grouping Customers with similar purchasing patterns (Unko Ads dikhate hai same!) Refail -> Retailers use custoring to segment customers based on demographics, purchase history or shopping preferences.
	21 Thage Segmentation:
	Medical Imaging -> MRI, CT Scan me clustering use kiya jaata hai to identify and analyse specific regions of interest
	· Satallite Imagery -> Clustering helps identify and classify objects or features in satellite images such as land cover types, regetation.
	3/ Recommendation Plustering:
	· lontont-based litering -> Q2.