Big Data To Avoid Weather Related Flight Delays

Pushkar Zagade

27 March 2015

A
Seminar
On
"**Big Data To Avoid Weather Related Flight Delays** "
Presented By
Mr.Pushkar Girish Zagade
Roll No. [TE-49]
Under the Guidance of
Prof. Neelima Satpute
Department Of Computer Engineering

JSPM's JAYAWANTRAO SAWANT COLLEGE OF ENGINEERING
Hadapsar Pune
Academic Year 2014-15

OUTLINES

- Introduction To Domain
- Identified Problem
- Motivation
- Purpose And Scope
- Objectives and planned Outcomes
- Literature Review (Existing Methods)
- Proposed System Components and Working Process
- System Features
- Mathematical Model and Analysis
- Algorithmic Survey to finalize the algorithm
- Survey of the selected Algorithm
- Results

OUTLINES (CONTINUED)

- ▶ Implementation technology
- ▶ Implementation using data tables
- ▶ Comparative outcome graphs
- ▶ Merits and Demerits
- ▶ Conclusion and Future Scope
- ▶ References
- ▶ Questions And Answers

INTRODUCTION TO DOMAIN

- This topic is generally belongs to weather forecasting that is how we implement Big Data computing for future weather prediction.
- In this lecture we will see how big data can be used to avoid weather Related Flight Delays.
- We will see Following Concepts in this session:
- **What Is Big Data ?**
- **How We Use Big Data Computing Method For Future Weather Prediction ?**
- **Data Mining in field of Bid Data To predict weather.**

IDENTIFIED SOCIAL PROBLEM

- As we know because of bad weather everyday lots of Flights has been canceled or dalayed.
- This is a big SOCIAL ISSUES we need to solve to avoid flight delays.
- As a Software Engineer , Engineers from University of Michigan developed a Big Data Computing method to predict the future weather .
- By using Big Data Computing method , they try to predict the future weather to avoid weather related flight delay.

MOTIVATIONS

- Organization Need Petabyte Of Memory To Store Large Amount Of Data.
- The Speed Of The Buisness For The Data Is Increasing Rapidly.

WHY THE DATA IS SO BIG?
They are now creating 2.5 quintillion bytes of data daily-Of the total data 90 percent of the data in the world today is created in last two years.

WHY WE NEED BIG DATA ?

- At present Social Media need large amount of data to store and analyze it.
- To give access to different users through out the world.
- To maintain security of data.
- For every domain , there should be large amount of data that need to be used and analyze for new development.

Problems With Big Data....

- The main problem with large data is to handle such large amount of data and it is very difficult to analyze .
- Processing it using conventional systems takes too long.
- It is very expansive.

Purpose And Scope:

Need Of Big Data For Weather Forecasting

- Weather is everything that affect our lives on a daily basis and has the power to make permamnent changes to a person's life or buisness.
- The Climate Corporation processes large weather data sets in excess of 50 terabytes of data.
- On a daily basis , the corporation utilizes 2.5 million weather measurements.

Objectives and planned Outcomes

- **The chief goal is to enable airlines to anticipate and deal with delays before they happen, says Amy Cohn, a U-M associate professor in industrial and operations engineering who researches airline industry operations.**

- **Big Data Management System Provides Weather Companies To Get Previous Decades Weather Trends And Patterns.**

- **By Analysing These Large Data Sets, Weather Predictions Are Becoming More Accurate And Being Quicker.**

Existing Method

- Prediction of Weather through Satellite is most commonly used method of weather prediction.
- Since it give appropriate result , but it predict the weather of some limited days about 2 weeks.
- Big Data Computing method overcome this limitation of Satellite system , So future weather become easy to predict.

Brief Working

- The students from the University of Michigan have started a new research which helps in understanding the weather of a particular place.

- They've gathered more than 10 years of hour-by-hour weather observations and domestic fight data, and are using advanced data analytics to spot patterns and help airlines manage more efficiently..

- The project uses public data that has been available for years, its size and scope make it unique, says Brian Lemay, a U-M doctoral student in industrial and operations engineering who leads the project.

Brief Working

- "We're the first to gather this data in one place and apply this level of computing power to it," Lemay said. "That enables us to do a very sophisticated analysis of how weather and flight delays are connected and go far beyond individual airports.

- The chief goal is to enable airlines to anticipate and deal with delays before they happen, says Amy Cohn, a U-M associate professor in industrial and operations engineering who researches airline industry operations.

- This is a very creative and new process. It could lead to understanding similarities in the weather in the past years. It could be of help in predicting the weather in the future.With the help of this data, the flights can be cautious of bad weather in advance.

Application

"Imagine you're scheduled to fly out of Detroit four hours from now and there's a storm in Atlanta." Cohn said. "The airline could use this data to determine that the storm in Atlanta is likely to delay your plane. They could then contact you and offer you a seat on an alternate flight. You save time, and the airline doesn't have to accommodate you on a later flight after the delay happens."

Data Mining In Big Data

- Generally, data mining is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both.

- Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified.

- Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

System Features

- **Large quantities of data**
  The volume of data so great it has to be analyzed by automated techniques e.g. satellite information, credit card transactions etc.

- **Complete data**
  This data is the characteristic of all data collection.

- **Easy data structure**
  Previous data are available which are complete so it become easy to analyse it.

Classification And Prediction Algrithm

- The previous data should be collected in order to analysed it.
- This algorithm consist of removal of unwanted data and selection of proper data.
- Apply Back Propagation Algorithm and classify data into collection and future prediction.
- About 80 percentage should be accurate by using this techniques.
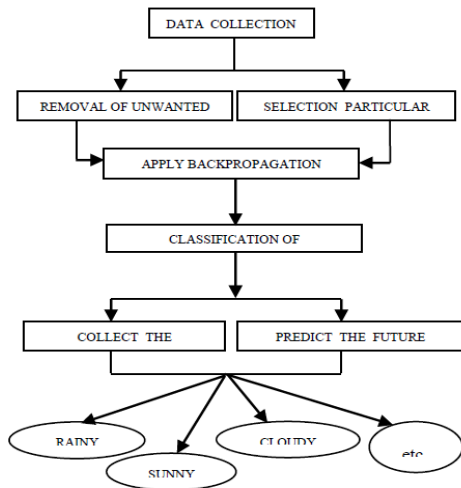
Data Flow Diagram

Fig 2: Data Flow Diagram

Survey Of Back Propagation Algorithm

- In order to perform a Back Propagation Algorithm a program or logic has to be created. What will be the change on other parameters by changing any one parameter, will be observed.

- After predicting that, what will be the weather in future after some period, the Classification method will take place.

- In Classification method, it will display what will be the future weather, whether it will be sunny day, rainy or cloudy day what will be the change in wind speed, humidity etc.

- The Classification Technique will help for taking some prevention from the climatic hazards.

Mathematical Model
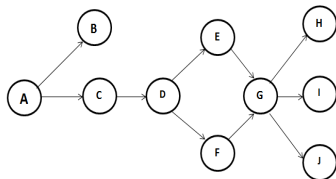
Where, A -.Collection Of Data
B - Waste Data
C - Useful Data
D - Apply BackPropagation Algo.
E - Prediction
F - Collection Of Information
G - Classification

Implementation Of Technoloy

STEP BY STEP DESCRIPTION OF OPERATION

- ▶ Data Collection
- ▶ Pre-processing
- ▶ Data Transfer
- ▶ Data Mining
- ▶ Prediction of Future Weather using ANN by Back Propagation Algorithm
- ▶ Classification

Example of DBMS, OLAP and Data Mining: Weather Data

DBMS:

Day	outlook	temperature	humidity	windy	play
1	sunny	85	85	false	no
2	sunny	80	90	true	no
3	overcast	83	86	false	yes
4	rainy	70	96	false	yes
5	rainy	68	80	false	yes
6	rainy	65	70	true	no
7	overcast	64	65	true	yes
8	sunny	72	95	false	no
9	sunny	69	70	false	yes
10	rainy	75	80	false	yes
11	sunny	75	70	true	yes
12	overcast	72	90	true	yes
13	overcast	81	75	false	yes
14	rainy	71	91	true	no

	2005	2006	2007	2008	2009	2010		AVE.	
January	28.4	39.8	32.2	27.9	21.3	23.3		28.8	
February	34.8	32.5	19.1	27.7	32.5	25.2		28.6	
March	37.6	41.1	47.3	39.3	45.7	44.4		42.6	
April	54.6	56.4	50.9	53.2	52.2	59.0		54.4	
May	60.6	61.9	68.0	59.5	63.0	65.5		63.1	
June	74.7	70.9	73.8	73.3	72.9	76.1		73.5	
July	76.1	77.7	73.8	74.9	70.6	77.4		75.1	
August	75.8	74.9	78.4	72.4	71.8	77.2		75.1	
September	70.1	64.3	70.5	68.9	66.8	68.3		68.2	
October	56.1	52.3	61.6	55.1	51.2	45.4		53.6	
November	45.1	45.2	42.5	40.9	46.5	43.4		43.9	
December	26.7	38.5	32.7	30.0	30.1	24.9		30.5	
Yearly Mean	53.4	54.6	54.2	51.9	52.1	52.4		53.1	
SPR (Apr-Jun)	63.3	63.1	64.2	62.0	62.7	66.5		63.6	
SUM (Jul-Sep)	74.0	72.3	74.2	72.1	69.7	74.3		72.8	
FAL (Oct-Dec)	42.6	45.3	45.6	42.0	42.6	37.9		42.7	
WIN (Jan-Mar)	33.6	37.8	32.9	31.6	33.2	31.0		33.3	

Comparative Outcome Graph

Merits
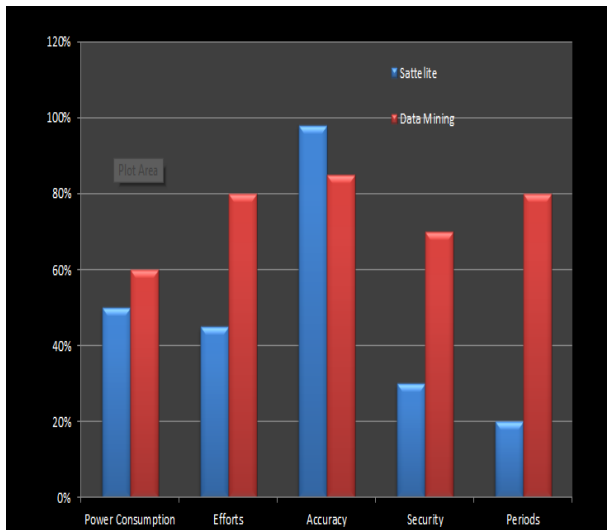
- Predict future trends, customer purchase habits
- Help with decision making
- Improve company revenue and lower costs
- Market basket analysis
- Fraud detection

De-Merits

- privacy/security
- Amount of data is overwhelming
- Great cost at implementation stage
- Possible misuse of information
- Possible in accuracy of data

Conclusion

**I think these new analytics will enable passengers as
well as airlines to better manage the whole travel
process. If airlines can offer more options and
passengers can educate themselves on how to use those
options, well see fewer delays and a less stressful travel
experience in the years to come.**

REFERENCES

1] http://www.csmonitor.com/USA/Society/2013/0811/The-new-age-of-algorithms-How-it-affects-the-way-we-live/(page)/3 viewed 9 Sep 2013

2] APA citation: How big data reduce weather-related flight delays (2014, December 23) retrieved 20 February 2015 from http://phys.org/news/2014-12-big-weather-related-flight.html

3] http://www.gartner.com/it-glossary/big-data/, viewed 15 Oct 2013.