**PIMPRI CHINCHWAD EDUCATION TRUST'S**

**PIMPRI CHINCHWAD COLLEGE OF ENGINEERING**



**DEPARTMENT OF FIRST YEAR ENGINEERING**

**A MINI PROJECT REPORT ON**

**“USE OF STATISTICS IN DATA MINING ”**

**DIVISION: A**

**SUBJECT: ENGINEERING MATHEMATICS III**

**SUBMITTED BY:**

1. **VIVEK BULANI – SECOA115**
2. **MOHIT BURKULE – SECOA116**
3. **PRAJAKTA JOSHI – SECOA116**
4. **BHAGYASHREE KADAM-**

**SUBMITTED TO: MR.ARIF BAGWAN SIR**

**DATE:**

**SIGN:**

**ACADEMIC YEAR 2018-19**

**SEMESTER I**

**INTRODUCTION**

**Statistics is the traditional field that deals with the quantification,**

**Collection, analysis, interpretation, and drawing conclusions from data.**

**Data mining is an interdisciplinary field that draws on computer sciences (data base, artificial intelligence, machine learning, graphical and**

**Visualization models), statistics and engineering (pattern recognition,**

**Neural networks).**

**Data mining involves the analysis of large existing data bases**

**In order to discover patterns and relationships in the data, and other**

**Findings (unexpected, surprising, and useful). Typically, it differs from**

**Traditional statistics on two issues: the size of the data set and the fact**

**That the data were initially collected for purpose other than the that**

**Of the Data mining analysis**

“Data mining is the process of exploration and analysis, by automatic or semiautomatic means, of large quantities of data in order to discover meaningful patterns and rules.”

(M. J. A. Berry and G. S. Linoff)

**STATISTICS AND DATA MINING**

**STATISTICS it is a process of understanding data and making certain decisions based on data whereas DATA MINING is the study of finding large patterns in data sets.**

**Hence data mining depends on many statistical techniques for its usage .**

**The Fourier transform is a generalization of the complex Fourier series in the limit of . L → ∞**

**F(s) = −∞ +∞∫ f ( x)e−2πisx dx**

**f ( x) =−∞ +∞ ∫ F(s)e 2πisx ds**

**Euler's formula : e ix = cos x + isin x**

**Since any function can be split into even and off portions:**

***f*E(x) = 1/2 [ f (x) + f (−x) ]**

***f*O(x) = 1/2 [ f (x) − f (−x) ]**

* **(x) = fE (x) + fO (x)**

**A Fourier transform can always be expressed in terms of the Fourier cosine transform and Fourier sine transform as**

**F(s)= −∞ +∞∫ fE (x)cos(2πisx)dx - *i* −∞ +∞∫ fO ( x)sin(2πsx)dx**

**This idea that a function could be broken down into its constituent frequencies (i.e., into sines and cosines of all frequencies was a powerful one and forms the backbone of the Fourier transform.**

**IMPORTANCE IN SIGNAL PROCESSING**

**A Fourier transform of a signal tells what frequencies are present in signal and in what proportions.**

**Example: Have you ever noticed that each of your phone's number buttons sounds different when you press during a call and that it sounds the same for every phone model? That's because they're each composed of two different sinusoids which can be used to uniquely identify the button. When you use your phone to punch in combinations to navigate a menu, the way that the other party knows what keys you pressed is by doing a Fourier transform of the input and looking at the frequencies present.**

**Apart from some very useful elementary properties which make the mathematics involved simple, some of the other reasons why it has such a widespread importance in signal processing are:**

**The magnitude square of the Fourier transform, |X(f)|^2 instantly tells us how much power the signal x(t) has at a particular frequency f.**

**The Parseval's theorem states that the total energy in a signal across all time is equal to the total energy in the transform across all frequencies. Thus, the transform is energy preserving.**

**By being able to split signals into their constituent frequencies, one can easily block out certain frequencies selectively by nullifying their contributions.**

**A shifted (delayed) signal in the time domain manifests as a phase change in the frequency domain. While this falls under the elementary property category, this is a widely used property in practice, especially in imaging and tomography applications Derivatives of signals (nth derivatives too) can be easily calculated using Fourier transforms.**

**APPLICATIONS IN SIGNAL PROCESSING**

**When processing signals, such as audio, radio waves, light waves, seismic waves, and even images, Fourier analysis can isolate narrowband components of a compound waveform, concentrating them for easier detection or removal. A large family of signal processing techniques consist of Fourier-transforming a signal, manipulating the Fourier-transformed data in a simple way, and reversing the transformation.**

**Some examples include:**

* **Digital radio reception without a superheterodyne circuit, as in a modern cell phone or radio scanner**
* **X-ray crystallography to reconstruct a crystal structure from its diffraction pattern**
* **Fourier transform ion cyclotron resonance mass spectrometry to determine the mass of ions from the frequency of cyclotron motion in a magnetic field**
* **Many other forms of spectroscopy, including infrared and nuclear magnetic resonance spectroscopies**
* **Generation of sound spectrograms used to analyze sounds**
* **Passive sonar used to classify targets based on machinery noise.**

**CONCLUSION**

**From this project we learnt the concept of series in signal processing,its uses and how the digital signal processing is widely used.**

**REFERENCE**

* ***Techtarget.com***
* ***Sciencedirect.com***
* ***Wikipedia.com***