iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** X **Month of publication:** October 2023

DOI: https://doi.org/10.22214/ijraset.2023.56095

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue X Oct 2023- Available at www.ijraset.com

Fraud Detection using IBMs Differential Privacy Library

Mrunmayee Waykar[1], Jasmine Joshi[2], Sanjyot Amritkar[3], Sanika Bharambe[4]
Department of Computer Engineering, Cummins College of Engineering for Women Pune, India

Abstract: ***In the year 2022, a total of 602 billion rupees were lost on bank frauds consisting of over 9103 major fraudulent cases. Due to the increasing number of frauds all around the world, it is vital to safeguard consumers' privacy. There is a prediction of a spike in cyber credit fraud in the future years. There are currently various techniques for credit card fraud detection but all the ML models have lower accuracies and cannot cope with the real-time datasets as the transaction data is extremely confidential. To address this issue we have used differential privacy for fraud detection. Differential privacy helps researchers to gain sensitive information about individuals without compromising their privacy. IBM provides a library: Diffprivlib, for exploring, researching and developing applications in Differential Privacy. In this paper, we have experimented with the impact of differential privacy on a sample dataset using various machine-learning models. With a comparative study of the algorithms, we propose that Isolation Forest has the highest accuracy. It can be added to the diffprivlib library since the library is open source and Isolation Forest does not exist in the library till date. Furthermore, we have built a system to predict fraudulent transactions.***
Keywords: Fraud Detection, Differential Privacy, diffprivlib, Isolation Forest

I. INTRODUCTION

Digital payment methods have gained popularity with increasing technological advancements. It had a lot of importance, especially during the Covid-19 period. Digital payments are advantageous to consumers since they simplify financial transactions. However, it also gave scammers more opportunities to exploit weaknesses and deceive customers in other ways. According to a survey, 42% of Indians experienced financial fraud in the previous three years, and 74% of those did not succeed in recovering their losses. With the help of differential privacy companies gather and share information about the user habits whilst safeguarding the individual's privacy. Diffprivlib is a general-purpose library for experimenting, investigating and developing applications in, differential privacy[8]. We, therefore, use IBM's differential privacy library, diffprivlib, which uses modules and algorithms to detect fraudulent transactions.

A. Understanding Differential Privacy

Differential privacy (DP) is a strong, mathematical definition of privacy in the context of statistical and machine learning analysis[6]. It prevents anyone from learning information about the individuals in a dataset by introducing a predetermined amount of randomness to the dataset. Differential privacy aims to ensure that regardless of whether an individual record is included in the data or not, a query on the data returns approximately the same result. DP provides a mathematically provable guarantee of privacy protection against a wide range of privacy attacks like differencing attacks, linkage attacks, and reconstruction attacks[11]. For example 911 calls : There are a number of calls that 911 operators receive every day. Each call provides a wealth of information about a person in need. This data can be used to see trends without compromising privacy. By slightly altering personal information we reduce the risk of breaching privacy while still allowing data to be shared. Altering is like adding a bit of noise to the data.
Differential privacy adds a privacy loss or privacy budget parameter to the dataset, which is frequently represented by the symbol epsilon (ε). It determines how much randomness or noise is injected into the underlying dataset. Differential privacy enables businesses to work with other organizations by sharing their data without jeopardizing the privacy of their customers.

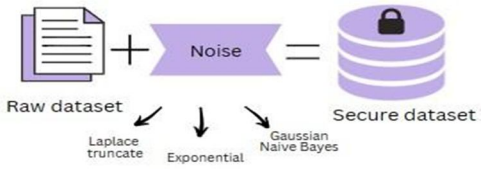


Fig 1.1 Differential privacy

B. Differential Privacy in Banking

Amongst all the various domains like healthcare, retail, banking, etc. privacy in banking has a crucial significance. These days banking companies are using more and more of our data via forms, ATMs, and online transactions to improve their services and have a clear database of our important yet private information like name, number, bank acc details, sensitive data, it can be risky because it can harm our privacy if it's leaked. Bank fraud is the illicit acquisition of funds from banks or other organizations using deceptive means. Several types of banking fraud include credit card fraud, phishing, forgery, identity theft, money laundering, loan fraud, etc. Due to this increasing number of frauds, fraud detection is the primary need of the hour.

IBM provides a library viz. diffprivlib, for exploring, researching and developing applications in Differential Privacy. Machine learning and differential privacy are the focus of this library. Its goal is to enable differentially private model experimentation, simulation, and implementation utilizing a shared codebase and building pieces. Python 3.4 is required to run Diffprivlib. Python was chosen because of its usability, abundance of machine learning features, and vibrant user community. Accessibility for academics and programmers with all levels of privacy competence, from beginners learning about differential privacy to more experienced users wishing to create custom applications, is a key component of diffprivlib. Thus we use IBM's differential privacy library, diffprivlib, to detect fraudulent transactions.

C. IBM's Differential Privacy Library "diffprivlib"

The goal of diffprivlib is to enable differentially private model experimentation, simulation, and implementation using a shared codebase and building pieces. The library has many tools that are the building blocks of differential privacy, and it also has applications for machine learning and data analytics. The developers focused on making the library easy to use and understandable, so it can be helpful for both beginners and experts in privacy. Anyone who wants to learn about data privacy or contribute to the development of models can use this library.

The library comprises four main parts:

1) *Mechanisms:* A group of mechanisms that serve as the foundation for applications that utilize differential privacy. Function-specific methods are used to converse with mechanisms, avoiding documentation and the need for duplicate code. Mechanisms include Laplace Truncate, Exponential, Binary, Bingham, Gaussian, etc.
2) *Models:* A set of machine learning models with varied levels of privacy. The models created by diffprivlib were designed to behave and have a syntax similar to the privacy-agnostic versions found in Scikit-learn.[2] For example, Classification models include Gaussian NB, Random Forest, Logistic Regression, etc. Clustering models include K-means.
3) *Tools:* A group of utilities and tools for basic data analytics with differentiated privacy. The module currently contains statistical methods for calculating the differentially private mean, variance, and standard deviation of a number array, as well as histogram algorithms for generating differentially private histograms on data. [2] For example, Histogram, Quantile like functions and general utilities.
4) *Accountant:* Privacy budget parameter for differential privacy. To keep track of privacy expenditures across queries and other data accesses, this class builds a privacy budget accountant. The accountant can be set up without a maximum budget, allowing users to track the full cost of their actions' privacy without any restrictions.

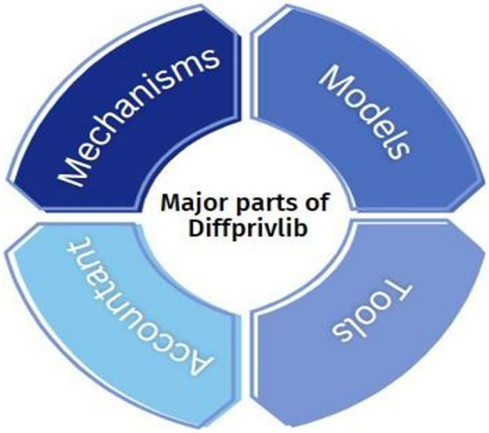


Fig 1.3 Major Parts Of Diffprivlib

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue X Oct 2023- Available at www.ijraset.com

II. FRAUD DETECTION

Banking fraud is when a third party uses unlawful means to gain access to your bank account. It involves the illegal acquisition of financial resources, sensitive information, or assets belonging to individuals, businesses, or the bank itself.

There are various types of banking fraud. Some of them include:

1) *Credit Card Frauds:* Credit card fraud happens when the fraudster obtains our credit card information and then uses it to make unauthorized transactions.
2) *Phishing:* It is a cybercrime where the individual gives out sensitive information via fraudulent phone calls, emails or website links.
3) *Identity theft:* This happens when a person's details like name, mobile number, ID number, address, bank details etc are stolen to carry out fraudulent activities.
4) *Cheque Frauds:* This involves forging or altering checks to withdraw money from the bank account.
5) *ATM Skimming:* Here the fraudster installs a device in the ATM which captures the card details and PINs to create a clone of the cards and withdraw money.

A. *Credit Card Frauds:*

Credit card frauds occur due to the illegal utilization of an individual's credit card information. It happens when an individual's personal credit card details are obtained without their permission and used to conduct unauthorized transactions or purchases.
Our main focus in this paper will be on credit card frauds and its detection.
Some types of credit card fraud include:

1) Counterfeit and Skimming frauds: When the fraudster copies data from the magnetic strip of the credit card and uses it to create a fake card.
2) Creating and using new credit card accounts under someone else's name
3) Dumpster diving: The criminal searches through the discarded items of the banking institutions and tries to find sensitive data in the same.
4) Manipulating the payer into making multiple credit card transactions against the same invoice/purchase

B. *Credit Card Fraud Detection*

There are currently various techniques for credit card fraud detection like Logistic Regression, Support Vector Machines, Naive Bayes, K-Nearest Neighbour etc. But all of the above ML models have lower accuracies and cannot cope with the real-time datasets as the transaction data is super confidential and cannot be used to train Machine learning models.
To address this issue we make use of Differential Privacy to train our model to maintain privacy and preserve accuracy.

The steps of fraud detection are as follows:

1) Data Pre-processing: Before using the library our data is gathered and then cleaned, normalized and prepared for the next steps.
2) Feature selection: This step involves engineering columns from the data that are relevant to training our model and dropping the ones that aren't.
3) Securing dataset: As mentioned above we have used IBM's Differential Privacy Library: Diffprivlib to secure our dataset. We do that by the various mechanisms present in the library like Laplace Truncated, Exponential, Binary etc to the sensitive fields. These mechanisms add some randomized values in the dataset or we call it adding noise before sending it to the model training as the customer information that we have is sensitive and we cannot directly use it for building the module.
4) Model training: Post-adding noise various models from the Diffprivlib library are tested on the data to train the model and see which gives the maximum accuracy and is the most optimized to give the best results. Some models include Gaussian NB, Random Forest, Logistic
5) Regression, K-means etc.
6) Anomaly Detection: The trained model is then used for anomaly detection on the real-time data in the system. When a new transaction happens, the model assigns a probability or score reflecting the possibility that the transaction is fraudulent. transactions that have unusually high probability or high scores fall outside the normal patterns and are flagged as suspicious. These transactions are probably fraudulent and are hence sent back for inspection.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue X Oct 2023- Available at www.ijraset.com

III.IMPLEMENTATION

In this section, we go through the work of IBM's diffprivlib library with its attributes like modules, models and tools in our system.

A. Dataset

We have worked with a sample dataset from Kaggle which contains columns like Account No, Transaction Amount, Average Transactions per day, and some boolean values as well like if the card has been declined if the card is foreign if it is from a high-risk country and so on. The first step is to clean the data which includes eliminating duplicate values, dealing with missing values, and converting the data into a suitable format for further study. As you can observe we have some sensitive fields which can breach a customer's privacy and lead to identity theft crimes. We add diffprivlib noise to these sensitive fields through modules like Laplace truncated and exponential to avoid that. All of these mechanisms help to ensure that the output of a function or data analysis process does not reveal sensitive information about the underlying data, while still providing useful and accurate results. These mechanisms are crucial for achieving differential privacy in a wide range of applications.

B. Model Training

After cleaning the data, it's important to divide it into two sets: the training set and the testing set. The training set is what we use to teach the machine how to analyze the data, and the testing set is how we check to see how well the machine can do it on its own. We usually divide the data so that 80% is used for training and 20% is used for testing, but this can change depending on the situation. There are many different types of machine learning models to choose from, including linear regression, decision trees, random forests, and neural networks, among others. The choice of model will depend on the dataset that we are trying to predict and the nature of our data. It includes various models for differentially private machine learning, including:

1) *Logistic Regression:* A model that analyses a binary dependent variable with a logistic function.
2) *Gaussian Naive Bayes:* A Bayesian model that applies the theorem of Bayes with a presumption that its features are independent and distributed normally.
3) *Random Forest:* An ensemble learning method that, while training, creates a large number of decision trees and delivers the class.

All of these models are modified to be differentially private using techniques such as adding noise to the training data or modifying the objective function to incorporate privacy guarantees. These modifications help to ensure that the models do not reveal sensitive information about the training data while still providing useful predictions.
We apply them to our training set and obtain the results as follows:

C. Model Comparison

	Metrics	Isolation Forest	Gaussian NB	Random Forest	Decision Tree	Logistic Regression
0	True Negatives	499	513	525	510	456
1	False Negatives	26	12	0	15	69
2	False Positives	7	27	60	35	29
3	True Positives	83	63	30	55	61
4	Accuracy	0.94	0.93	0.90	0.91	0.84
5	Precision	0.76	0.84	1	0.78	0.46
6	Recall	0.92	0.7	0.33	0.61	0.67
7	F1- score	0.83	0.76	0.5	0.68	0.55
8	Support	None	None	None	None	None

After a thorough comparative study of various models that can be used for detecting frauds, we observe that the Isolation Forest algorithm has the maximum accuracy i.e. 0.943089.

IV. ISOLATION FOREST

Isolation Forest can outperform other anomaly detection algorithms and provide better accuracy in identifying outliers. The algorithm can be especially effective in scenarios where the dataset has a high dimensionality, contains noisy or irrelevant features, or has a large number of data points. One of the advantages of Isolation Forest is that it has a low computational cost and can handle very large datasets efficiently. The algorithm also does not require a priori knowledge of the normal or anomalous behaviour in the dataset, making it suitable for unsupervised anomaly detection tasks.

$$c(N) = 2H(N-1) - \left(\frac{2(N-1)}{N}\right),$$

where n is the testing dataset size, m is the sample set size and H is the harmonic number calculated by ln(i) + 0.5772, also known as the Euler-Mascheroni constant. Overall, Isolation forest can be a useful tool for detecting outliers in certain scenarios and can potentially provide better accuracy compared to other methods. Therefore, we propose to include the Isolation Forest model in IBM's diffprivlib library due to its maximum accuracy.

We applied Isolation Forest to our dataset and got the following results:
Test Accuracy: 0.943089

	precision	recall	f1-score	support
Not Fraud	0.99	0.95	0.97	525
Fraud	0.75	0.93	0.83	90
accuracy			0.94	615
macro avg	0.87	0.94	0.9	615
weighted avg	0.95	0.94	0.95	615

Table 4.1 Isolation Forest

The basic idea is to build a large number of fully random binary trees (randomly choose a feature, splitting threshold until each data point is in its leaf). The depth at which a point becomes completely isolated is the statistical measure we're most interested in here. The intuition here is that an outlier can be isolated with a few random splits, whereas a nominal point requires many splits (due to heavy density).

A Few of the significant advantages of Isolation Forest are :

1) It has reduced computational times due to the early and rapid detection of anomalies.
2) It is Scalable to high-dimensional and large-scale datasets.
3) Isolation Forest has a "low constant and low memory requirement," which translates to low overhead.

Therefore, we propose to include the Isolation Forest model in IBM's diffprivlib library due to its maximum accuracy.

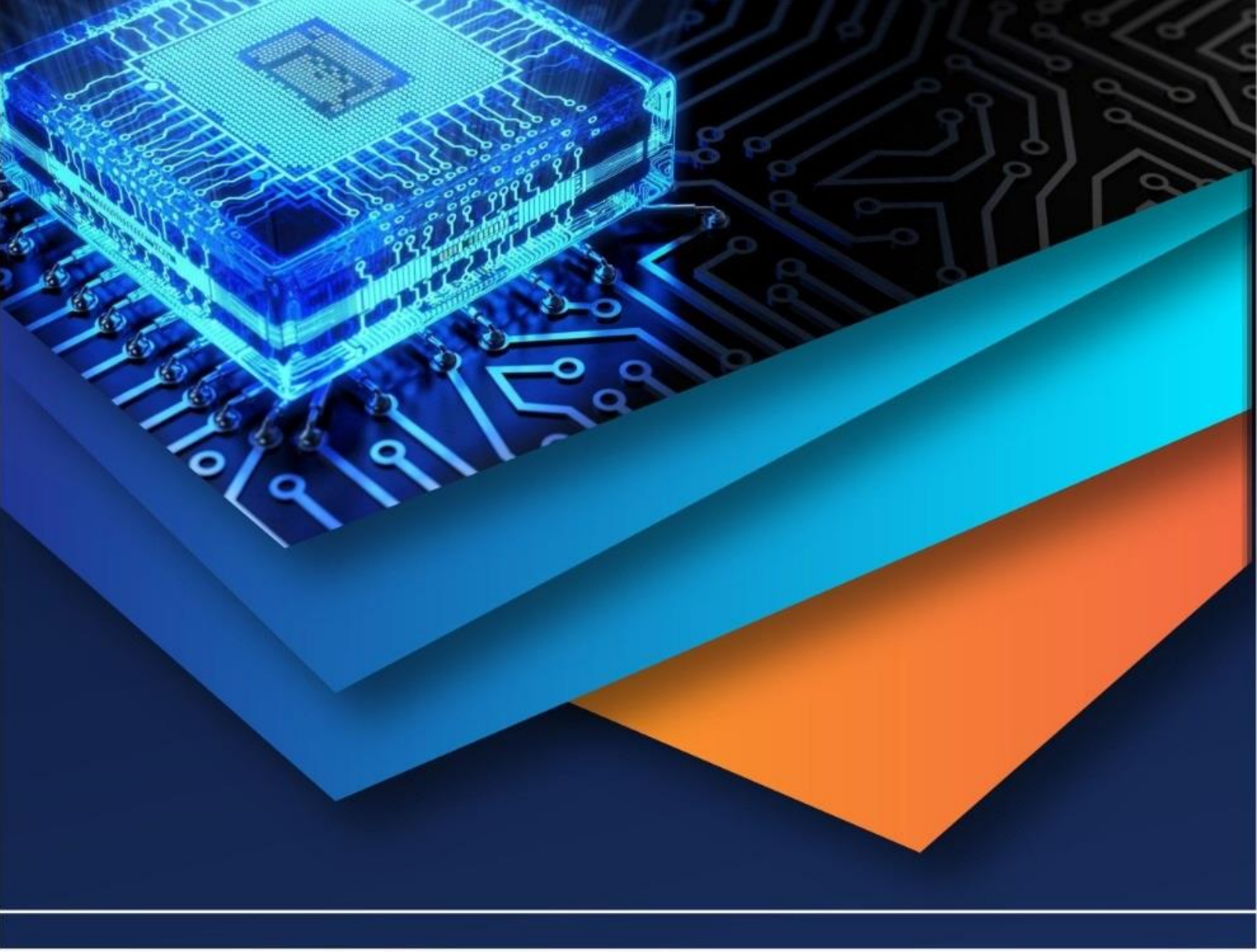
V. CONCLUSIONS

In this research, we have built a credit card fraud detection system, with the main aim being the utilization of differential privacy. Due to the increasing number of frauds all around the world, it is vital to safeguard consumers' privacy. Banks collect data to enhance their services, but we as customers care about our privacy and don't want our personal information to be at risk. To find a balance between these opposing desires, banks can use a technique called differential privacy. From the comparison of the machine learning models, we can conclude that the Isolation forest has the maximum accuracy. This algorithm does not exist in IBM's Differential privacy library "diffprivlib" to date. Hence, we intend to contribute to the library by adding the Isolation Forest model which can be further used for fraud detection.

Several avenues for further work are promising. In particular, we intend to work on real-time datasets to make our system more accurate and compatible with the complications of real-time data. Furthermore, we propose the idea of using differential privacy in not just credit card fraud detection but explore all other financial transactions. We also look forward to applying this algorithm to varied domains in future including healthcare, retail etc.

REFERENCES

[1] Maniar, Tabish & Akkinepally, Alekhya & Sharma, Anantha: Differential Privacy for Credit Risk Model, 2022
[2] Ileberi, Emmanuel & Sun, Yanxia & Wang, Zenghui. (2022). A machine learning based credit card fraud detection using the GA algorithm for feature selection. Journal of Big Data. 9.10.1186/s40537-022-00573-8.
[3] Detecting Credit Card Fraud using Machine Learning, December 2021, International Journal of Interactive Mobile Technologies (iJIM) 15(24):108-122 , DOI:10.3991/ijim.v15i24.27355.
[4] R. Sailusha, V. Gnaneswar, R. Ramesh and G. R. Rao: Credit Card Fraud Detection Using Machine Learning, 2020 4th International Conference on Intelligent Computing and Control Systems.
[5] Holohan, Naoise & Braghin, Stefano & Aonghusa, Pol & Levacher, Killian. (2019): Diffprivlib: The IBM Differential Privacy Library.
[6] Kobbi Nissim , Thomas Steinke, Alexandra Wood, Micah Altman: Differential Privacy: A Primer for a Non-technical Audience, School of Engineering and Applied Science, Harvard University, February 2018.
[7] Diffprivlib mechanisms: https://diffprivlib.readthedocs.io/en/latest/modules/m echanisms.html#diffprivlib.mechanisms.Laplace
[8] Random Forest: https://www.ibm.com/topics/random-forest
[9] Gaussian Naive Bayes: https://iq.opengenus.org/gaussian-naive-bayes/
[10] Confusion Matrix: https://plat.ai/blog/confusion-matrix-in-machine-learn ing/
[11] Differential privacy in data centric organization: https://analyticsindiamag.com/why-is-differential-pri vacy-important-for-data-organisation/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY