



Initiatives on AI/ML based Analytics

Government eProcurement System of NIC, GePNIC[©] Government of India

African Public Procurement Network Webinar

July 2023

(Virtual Conference)

Background

National Informatics Centre, MeitY has developed e-Procurement software system, GePNIC[©] as a product for facilitating Electronic Procurement across Government.

This has been developed in consultation with Procurement Policy Division, DoE, Ministry of Finance, various State Governments and Central Public Sector Enterprises.

01

Implemented across

- 31 States & Union Territories
- 700+ Central Government Procuring Entities

02

Statistics for June 2023

- # e-Tenders : 0.17 Million
- Worth : \$ 36.47 Billion

03

eAuction India implemented across

- 21 States / UT's
- Covering 44,398 (Multilot and Single Lot Auctions)
- Worth \$1.91 Billion

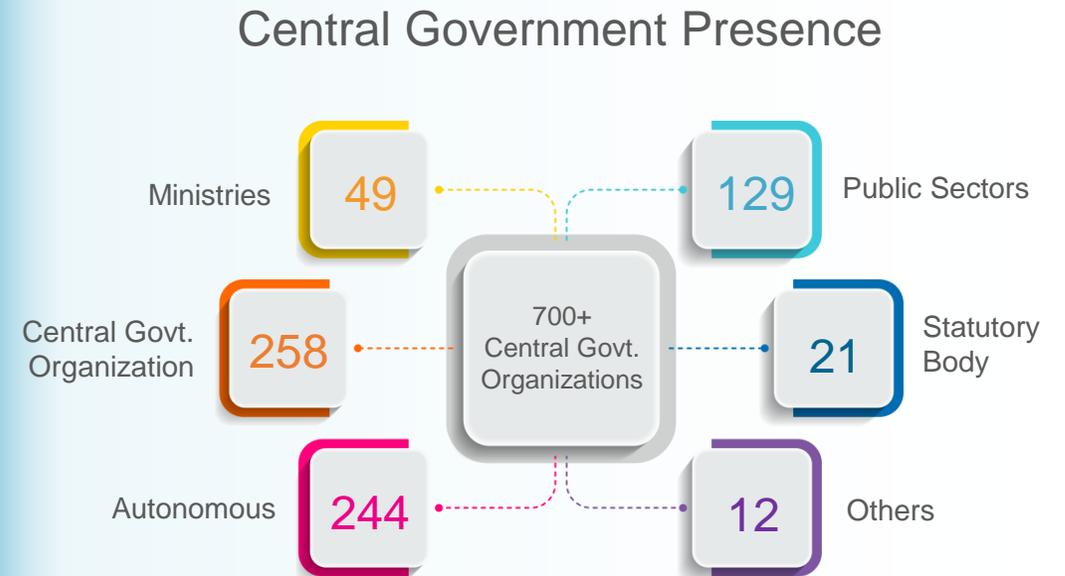
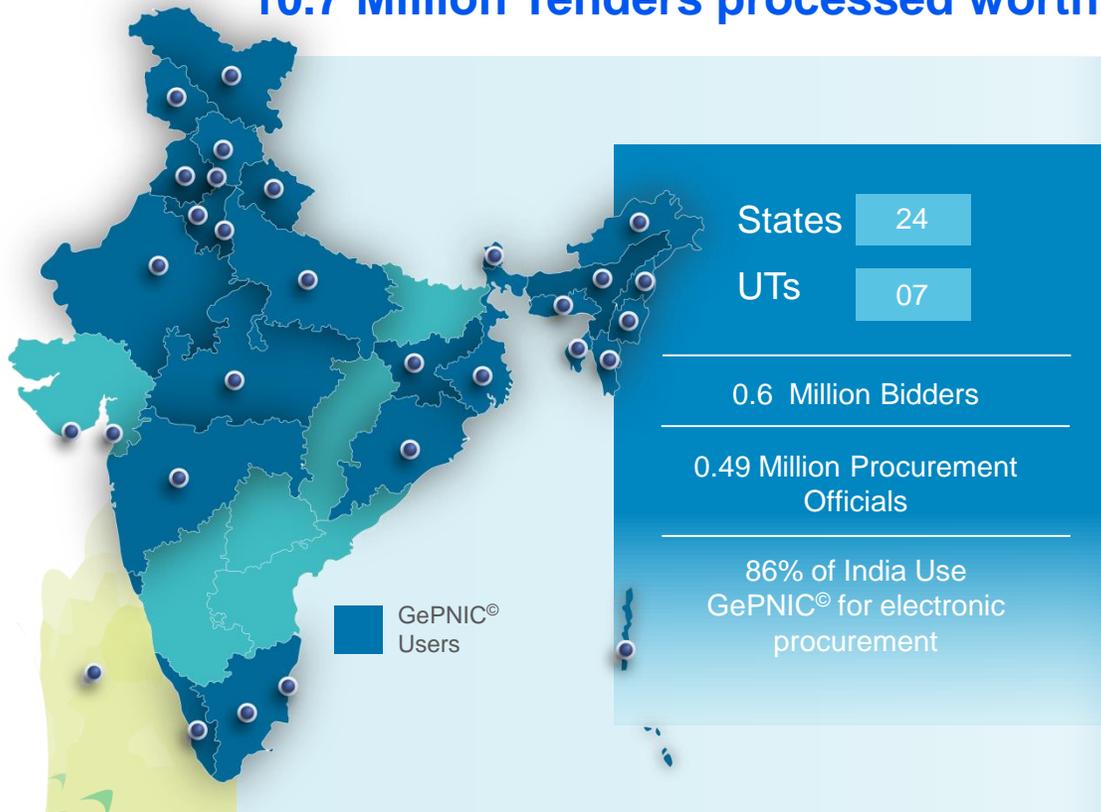
04

Cumulative Statistics since 2008

- # e-Tenders : 10.7 Million
- Worth : \$ 1951 Billion

GePNIC[©] at a Glance

10.7 Million Tenders processed worth \$ 1951 Billion , since 2008 with secure Bidding



Spans over 45 Live instances

Objectives of the System



Features of GePNIC[©] eProcurement System

Core Functional



- Single product with configurable workflow
- Caters to Goods, Works & Services
- Facilitates - RFP/RFQ/EOI/Global tenders with Multi Currency/Rate Contract
- Open/Limited/Single/Tender cum Auction/QCBS
- BOT/BOOT/Turn-key Projects/Empanelment/EPC Contract/Multi-stage/Piece-work/PPP-Bot-HAM
- Customizable Price Bid templates-BOQ to suit Turnkey and Public Private Partnership projects
- Adheres to GFR 2017, MDB guidelines, State and CPSE Rules
- Centralized Debarment information
- System Malfunctioning and recovery process
- Shortfall Documents during Technical Evaluation
- Supports bulky Bid documents, size not a deterrent
- Bid Acknowledgement receipt for Bidder after bid submission
- SMS-Mail alerts/Mobile App/Dashboard/Chat Bot
- Real time Dashboard with Key Parameter Indicators

Features of GePNIC[©] eProcurement System

Configurable



- Tender cum Auction-Before/After Financial Opening
- eReverse Auction with or without H1 elimination
- Item Wise Evaluation
- General Technical Evaluation with Comparative Chart
- Critical Date Validations
- Visibility of Bid Documents to Participating Bidders
- Two Stage Bidding
- Non Disclosure Agreement

- Tender Revocation to Technical/Financial Opening Stage
- Pre and Post Tender Integration
- Auto Tendering Process with Technical Parameter Sheet
- Preferential Bidder treatment for MSME, Startup etc.
- Auto extension on pre specified number of Bids
- Holiday/Calendar Mapping
- Workflow based technical evaluations
- Finance Rebid Submission for Negotiation/Tie-Break

Features of GePNIC[©] eProcurement System

Security



- Two factor authentication using Digital Signature Certificate (DSC)
- Bid encryption by DSC of Dept. user
- Encryption of Technical & Financial bids at client end using PKI Technology
- Digital Signing of all documents
- Log Shipping, Network Time Protocol (NTP) Configuration
- Secured Hosting in ISO 27001: 2013. Certified Data Centre
- Integration with external Systems
- Periodic audits by STQC as per E-Procurement System (EPS) Guidelines of MeitY
- Replication/Backup with Retention Lock
- 24X7 availability of the system
- Certified by World Bank, ADB
- Business Continuity and Disaster Recovery (DR)
- Active - Active DR Setup
- Continuous Integration/Continuous Delivery (CI/CD)
- Transparency through Tender Status in Public Domain

Global Tendering using GePNIC[©]

01 Facilitates Multi Currency Global Tenders

02 Customisable Bill of Quantity - BOQ

03 Digital Signature Certificates – Foreign Citizens

04 Used by CPSEs, Central Govt and Few States.

05 Widely used in Multilateral Bank Tenders like ADB, JAICA, World Bank etc.

06 An average of 2500 Global Tender per month worth Approx. Rs. 590,800 Millions.



Introduction of AI/ML in GePNIC

Initiatives to carryout Artificial Intelligence/Machine Learning (AI/ML) based analytics since June 2019.

Few select areas were identified for analysis, as good volume of relevant data is readily available

To start with, a sample set was restricted to 50,000 records for building the model

This model was then applied to data related to tenders published during the period from April 2016 to March 2019 for analysis

Focus Areas for AI/ML analytics



01

Bidder Segmentation Analysis - based on tender value for which the bidders have participated - to provide focused services to bidders

Analysis on pattern of association - Bidder Participation trend analysis for pre-emptive Alerts

02

03

Forecasting of Volume of Transactions - to envisage and provide Compute and Storage requirements

Invalid Names - Automatic detection of invalid names entered during user registration process

04

Focus Areas for AI/ML analytics



01

Bidder Segmentation Analysis - based on tender value for which the bidders have participated - to provide focused services to bidders

Analysis on pattern of association - Bidder Participation trend analysis for pre-emptive Alerts

02

03

Forecasting of Volume of Transactions - to envisage and provide Compute and Storage requirements

Invalid Names - Automatic detection of invalid names entered during user registration process

04

Bidder Segmentation Analysis



- a) Type of tenders they participate often
- b) Products for which they are bidding
- c) Category of Tenders
- d) Value of tenders

Helps us to segment each bidder to a particular category and this input can be used to for further analysis like send specific Alerts , Predicting participation %

By segmenting each bidder to a particular category, focused services can be planned

Clustering Method is applied to group the Bidders in these categories

K-Prototype Clustering



It belongs to a group of **unsupervised learning** algorithm

Clustering is the task of grouping a set of objects in such a way that objects in the same cluster are more similar to each other than to those in other groups

k-prototypes clustering aims to partition n **observations** into k **clusters** in which each observation belongs to the cluster with the nearest centroid.

Elbow Method is used to find the optimal number of clusters.

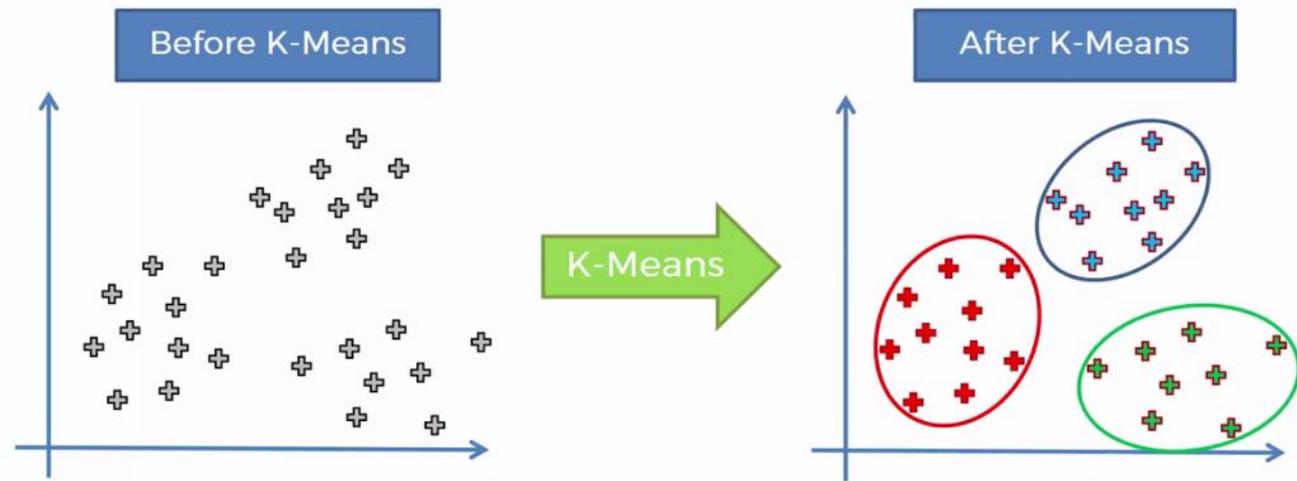
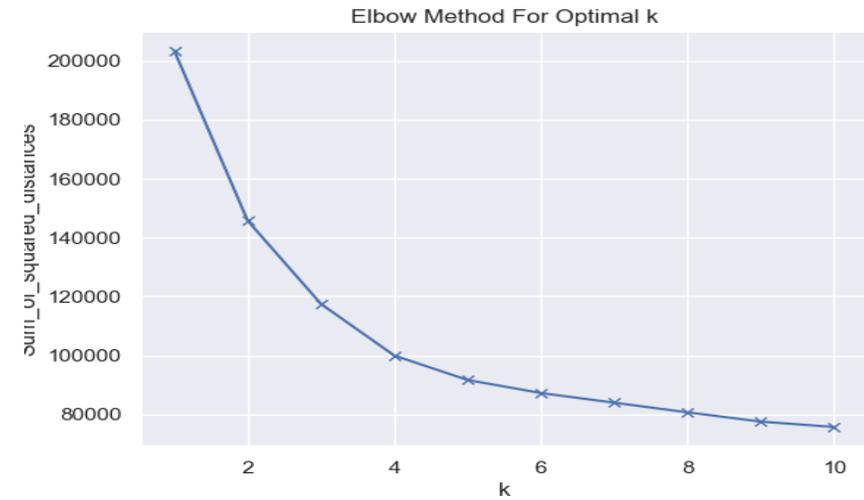
Visualisation of Data Clusters



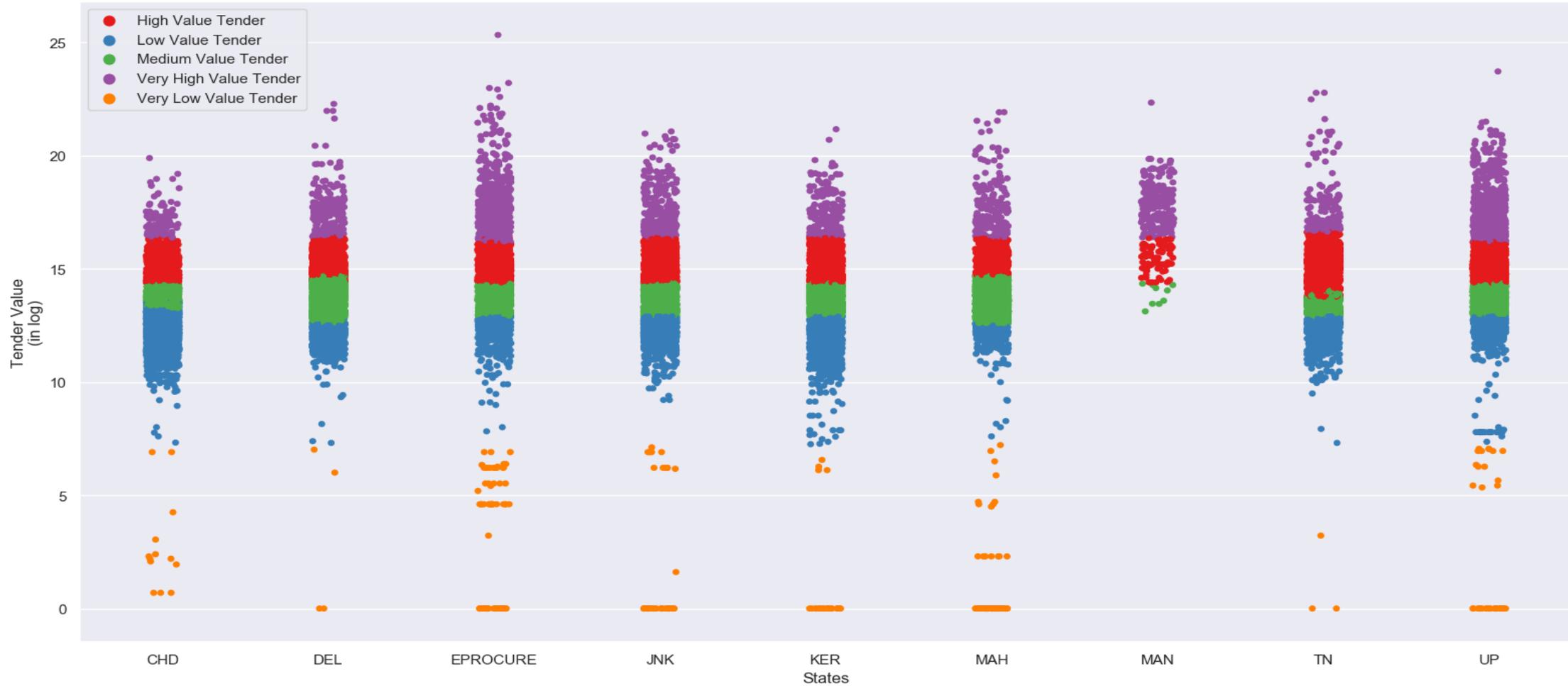
An Elbow shape is formed at the 5th cluster point. So, we choose $k=5$ as our optimal number of clusters.

The 5 bidder clusters formed can be interpreted as :

- Very low tender value
- Low tender value
- Medium tender value
- High tender value
- Very high value



Segmentation of Bidders amongst the various provinces/ implementations



Inference drawn from these Segmentation



Inference from the previous graph indicates a clear picture on participation pattern of Bidders in various Provinces - Chandigarh, Delhi, eProcure, J&K, Kerala, Maharashtra, Manipur, Tamil Nadu, UP are given below [applied For Tender Value]

“ *High Value Tenders are more in Tamil Nadu State and Medium Value Tenders are less* ”

“ *Very High Value tenders are more in Central Govt & Uttar Pradesh* ”

“ *Very Low value tenders are sporadic in all States* ”

“ *Those bidders who participate in High Value tenders do not participate in low value. (Inference)* ”

This can be applied to detect the pattern on – Last Hour Bidding , Re-Submissions , Effective Time taken to Bid

Focus Areas for AI/ML analytics



01

Bidder Segmentation Analysis - based on tender value for which the bidders have participated - to provide focused services to bidders

Analysis on pattern of association - Bidder Participation trend analysis for pre-emptive Alerts

02

03

Forecasting of Volume of Transactions - to envisage and provide Compute and Storage requirements

Invalid Names - Automatic detection of invalid names entered during user registration process

04

Analysis on pattern of Association



01

Various platforms like Amazon, Flipkart, Walmart **recommends products** based on the **user behavior**



The model is capable of automatically detecting and **recommending tenders** to the bidders based on the **past behavior** of the bidders.

02

03

A **push notification** about such new tender will be sent to the bidders who are highly likely to participate in that.



Using this recommendation, the number of bidders participating in a tender can be **increased significantly** as the right bidders are targeted.

04

05

It also ensures timely automatic alerts to the bidders



APRIORI



The *Apriori* Algorithm was proposed by Agrawal and Srikant in 1994.



Apriori is an association rule mining algorithm used to identify frequent item sets.



The three important metrics of apriori are support, confidence and lift.

- **Support** - Percentage of a categories that contains the item or item set.
- **Confidence** - Percentage of times that category B is bidden, given that category A was bidden
- **Lift** - Measure for the relationship between A and B



PRODUCT CATEGORY ASSOCIATION



antecedents	consequents	support	confidence	lift
frozenset({'Computer- Data Processing', 'Miscellaneous Works', 'Computer- H/W'})	frozenset({'Miscellaneous', 'Civil Works - Construction Works', 'Electronics Equipment'})	0.0414707	0.889908	12.2441
frozenset({'Computer- Data Processing', 'Miscellaneous Works', 'Miscellaneous'})	frozenset({'Civil Works - Construction Works', 'Computer- H/W', 'Electronics Equipment'})	0.0414707	0.873874	11.7471
frozenset({'Computer- S/W', 'Miscellaneous Works', 'Civil Works - Construction Works'})	frozenset({'Miscellaneous', 'Computer- H/W', 'Electronics Equipment'})	0.0427533	0.917431	11.7261
frozenset({'Computer- S/W', 'Civil Works - Construction Works', 'Food Products'})	frozenset({'Miscellaneous', 'Computer- H/W', 'Electronics Equipment'})	0.0401881	0.903846	11.5524
frozenset({'Computer- S/W', 'Stationery', 'Civil Works - Construction Works'})	frozenset({'Miscellaneous', 'Computer- H/W', 'Electronics Equipment'})	0.0401881	0.903846	11.5524
frozenset({'Computer- S/W', 'Miscellaneous', 'Food Products'})	frozenset({'Civil Works - Construction Works', 'Computer- H/W', 'Electronics Equipment'})	0.0401881	0.854545	11.4873
frozenset({'Computer- Data Processing', 'Miscellaneous Works', 'Civil Works - Construction Works'})	frozenset({'Miscellaneous', 'Computer- H/W', 'Electronics Equipment'})	0.0414707	0.873874	11.1693
frozenset({'Computer- S/W', 'Food Products'})	frozenset({'Miscellaneous', 'Computer- H/W', 'Electronics Equipment'})	0.0440359	0.872881	11.1567
frozenset({'Computer- S/W', 'Stationery'})	frozenset({'Miscellaneous', 'Computer- H/W', 'Electronics Equipment'})	0.0436084	0.871795	11.1428
frozenset({'Miscellaneous Works', 'Civil Works - Construction Works', 'Food Products'})	frozenset({'Miscellaneous', 'Computer- H/W', 'Electronics Equipment'})	0.0431808	0.87069	11.1287
frozenset({'Computer- Data Processing', 'Miscellaneous Works', 'Miscellaneous', 'Computer- H/W'})	frozenset({'Civil Works - Construction Works', 'Electronics Equipment'})	0.0414707	0.92381	11.081
frozenset({'Medicines', 'Food Products'})	frozenset({'Miscellaneous', 'Computer- H/W', 'Electronics Equipment'})	0.0410432	0.857143	10.9555
frozenset({'Stationery', 'Food Products'})	frozenset({'Miscellaneous', 'Computer- H/W', 'Electronics Equipment'})	0.0410432	0.857143	10.9555

It can be inferred that the higher the 'lift' value, the more the chance of association. i.e. Those bidders who have bid for tenders for 'Computer Data Processing' would also bid for tenders for 'Computer-HW'.

Focus Areas for AI/ML analytics



01

Bidder Segmentation Analysis - based on tender value for which the bidders have participated - to provide focused services to bidders

Analysis on pattern of association - Bidder Participation trend analysis for pre-emptive Alerts

02

03

Forecasting of Volume of Transactions - to envisage and provide Compute and Storage requirements

Invalid Names - Automatic detection of invalid names entered during user registration process

04

Forecasting Analytics



SARIMA=Seasonal + AutoRegression + Integration + Moving Average

To use SARIMA there are three steps

Define the model

Fit the defined model

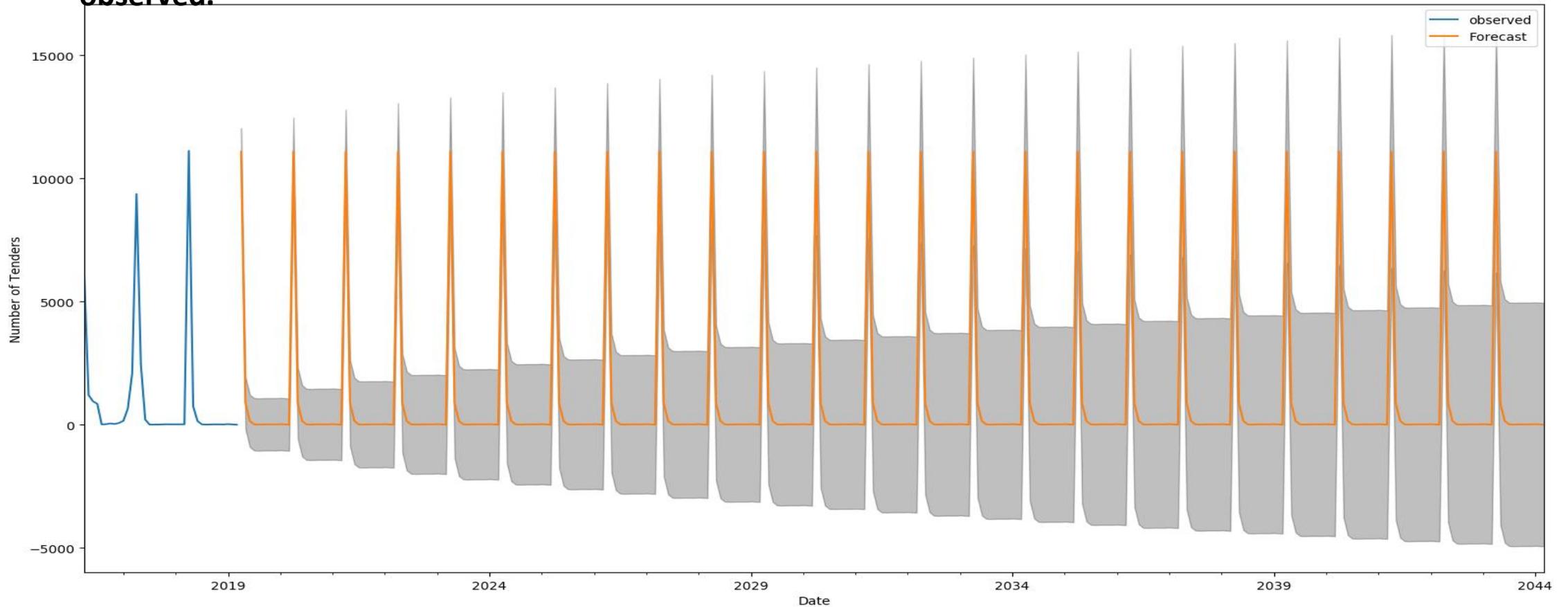
Make a prediction with the fit model

The model can be used to forecast the number of tenders to be handled, Peak time Bids , Peak Month load in future in GePNIC eProcurement platform

A) Forecast using SARIMA model



Based on the three years data, the forecast for the next few years have been carried out. It may be noticed that the years in which the actuals are almost in line with the forecasted value can be observed.



Focus Areas for AI/ML analytics



01

Bidder Segmentation Analysis - based on tender value for which the bidders have participated - to provide focused services to bidders

Analysis on pattern of association - Bidder Participation trend analysis for pre-emptive Alerts

02

03

Forecasting of Volume of Transactions - to envisage and provide Compute and Storage requirements

Invalid Names - Automatic detection of invalid names entered during user registration process

04

Auto Detection of Invalid Names



To flag the users who have entered invalid names, an intelligent model is created to classify the names.



Real names:-Indian and Foreign Names



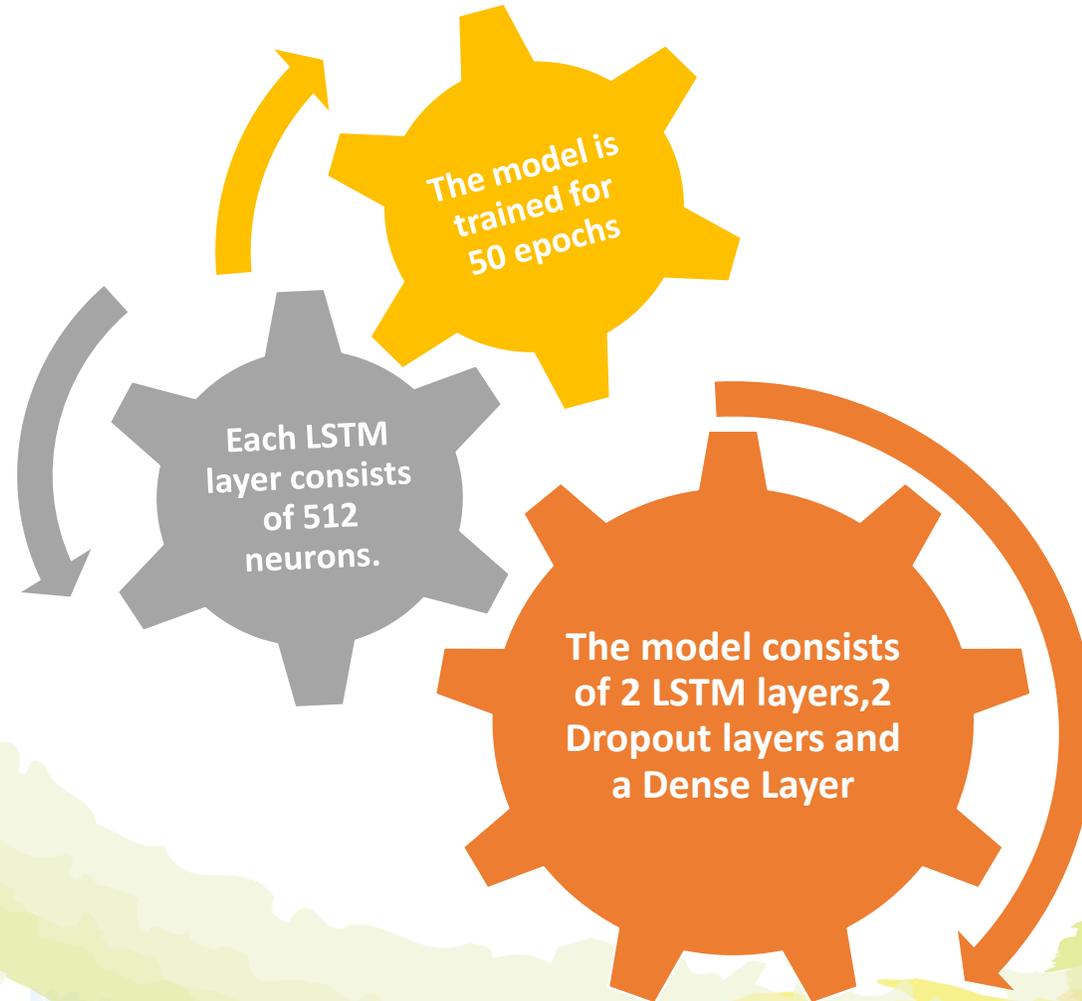
Invalid names:-common verbs, adjectives, names generated from a random word generator.



The size of the total dataset is 2 lakh names



Long Short-Term Memory LSTM Model Parameters



Model's Predictions



```
Type the Name  
krishna kanth  
krishna kanth is a Original Name
```

```
Type the Name  
rupam  
rupam is a Original Name
```

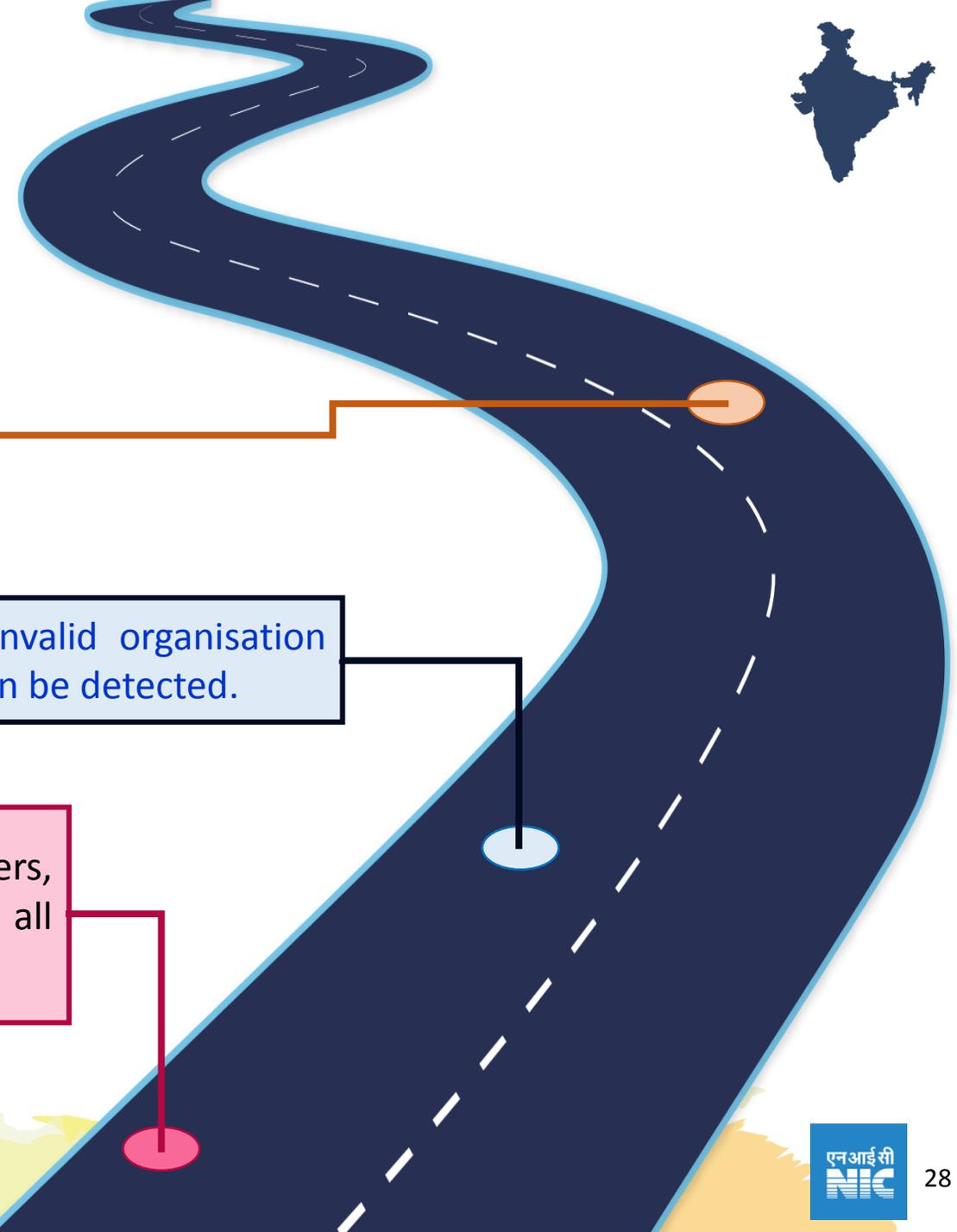
```
Type the Name  
priya  
priya is a Original Name
```

```
Type the Name  
gem,m gejkjgnjgn  
gem,m gejkjgnjgn is a Invalid Name
```

```
Type the Name  
beautiful  
beautiful is a Invalid Name
```

```
Type the Name  
nobody  
nobody is a Invalid Name
```

Road Ahead



Various other quantitative measures like number of bidders , number of users, number of corrigendum etc in a particular month can be predicted similar to the number of tender's method.

Similar to the Invalid User Names, Invalid organisation names , product category names etc can be detected.

We can also use other models to identify fake bidders, Bidders who participate and get 100% bid winning in all tenders that they participate etc.,

Tools and Libraries used



- **Fbprophet-0.5**
- **Keras-gpu-2.2.**
- **Matplotlib-3.1.0**
- **Numpy-1.16.4**
- **Pandas-0.24.2**
- **Scikit-Learn-0.21.2**
- **Seaborn-0.9.0**
- **Statsmodels-0.9.0**

Important Links

Important Links

eGP Platform	https://eprocure.gov.in/cppp
Dashboard	https://eprocure.gov.in/eprocdashboard/
GePNIC [©] Product Information	https://gepnic.gov.in
eAuction India	https://eauction.gov.in
ADB/World Bank Certification	https://eprocure.gov.in/cppp/bank_docs
Awards to Stakeholders	https://gepnic.gov.in/show_content.php?lang=1&level=1&sublinkid=1790&lid=1707
Stakeholders Workshops	https://gepnic.gov.in/show_content.php?layout=&lang=1&level=1&sublinkid=828&lid=834
Monthly News Letters	https://gepnic.gov.in/newsletters.php?lang=1&EncHid=&lid=250
Testimonial(User Speak)-Sample	https://gepnic.gov.in/index.php#testimonial
Annual Reports	https://eprocure.gov.in/cppp/sites/default/files/eproc/Annualreports/CPPP-AnnualReport-2019-20.pdf
Awards & Recognitions	https://gepnic.gov.in/show_content.php?layout=&lang=1&level=1&sublinkid=1757&lid=1692
Security Quality Certification	https://eprocure.gov.in/cppp/sites/default/files/eproc/certificate/EPS_STQC_Certificate2020.pdf
eGP Bags TOP Rank in NeSDA	https://twitter.com/NICMeity/status/1537342838543118336?cxt=HHwWgMDT1dPK3tUqAAAA



Recognitions for GePNIC[®]/CPPP

Madhya Pradesh eTenders Portal bagged Technology Sabha Award under IoT category, in August 2022

CPPP secured Top Rank in NeSDA 2021 Assessment Report under Central Ministry Services Portal

Computer society of India 19th CSI SIG eGovernance Awards of Appreciation 2021 to MP Tender Portal

“PM Mementoes – Auction” conferred with Award of Appreciation in the 18th CSI SIG eGovernance Awards 2020

Coal India Limited conferred with an Award for automated e-tendering system for excellence in “Innovation in Procurement Process” at NATCOM 2019.

GePNIC[®] bagged Data Quality Challenge Award 2020 conducted by NIC.

Central Public Procurement Portal (CPPP) bagged the Award of Excellence 2020 in in the 18th CSI SIG eGovernance Awards.

GePNIC[®] won South Asian Procurement Innovation Award 2018 organised by World Bank and Procurement iNET

GePNIC[®] and Govt. of Kerala both won CSI Nihilent eGovernance Award 2017 and 2013 under best eGovernance Project

CPP Portal won Gems of India Award 2017.

GePNIC[®] won Two SKOCH Award for Smart Governance in 2015 and also in 2010

GePNIC[®] bags eIndia award 2014 for Govt. of Kerala and Mahanadi Coalfields Limited - Coal India Limited.

Govt. of Maharashtra awarded "Best Government to Business (G2B) Initiative of the Year 2013" during eMaharashtra Awards 2013

Mahanadi Coalfields Ltd. – Coal India Limited won CII-IT Award 2011

Odisha eProcurement implementation received eIndia Award 2009 "G2B initiation of the year"



Thank
you

Queries may be sent to
ks.raghavan@nic.in