

# **MINERAL TARGETING USING ARTIFICIAL INTELLIGENCE**

**HACKATHON 2025**

**TEAM GSI**

**TEAM MEMBERS:**

**DR. SAUDAMNI SAHOO  
NIHARIKA DAS  
PIU DHIBAR**

**Email ID: [saudamini.sahoo3@gmail.com](mailto:saudamini.sahoo3@gmail.com)  
[saudamini.sahoo@gsi.gov.in](mailto:saudamini.sahoo@gsi.gov.in)**

## TABLE OF CONTENTS

<b>INTRODUCTION .....</b>	<b>1</b>
Regional Geology .....	1
Geology of Karnataka .....	1
Geology of Andhra Pradesh.....	1
<b>OBJECTIVES .....</b>	<b>3</b>
<b>NAME AND DETAILS OF PARTICIPANT AND COMPANY .....</b>	<b>3</b>
Team Members: .....	3
Organisation:.....	3
Mail id.....	3
<b>HARDWARE AND SOFTWARE USED .....</b>	<b>3</b>
<b>METHODOLOGY: .....</b>	<b>3</b>
1. Data Collection: .....	3
2. Data Preparation: .....	4
3.Data Cleaning.....	5
<b>BRIEF DESCRIPTION OF THE AVAILABLE DATA .....</b>	<b>5</b>
50K geological Data .....	5
1. lithology_gcs_ngdr .....	5
2. Commodity shape files .....	6
25K geological Data .....	7
Geochemical Data Analysis and Interpretation: .....	8
1. Major Oxides (%): .....	8
2. Trace Elements (ppm):.....	8
3. Precious Metals (ppb): .....	9
Principal Component Analysis of Geochemical Data: .....	9
Interpretation.....	10
Anomaly Plot for Gold: .....	10
Anomaly Plot for Iron Oxide:.....	11
Comparison of Au and Fe <sub>2</sub> O <sub>3</sub> Map:.....	12
Analysis of Geophysical Data.....	13
Introduction.....	13
Geophysical Data processing and preparation of maps .....	13
Results and Discussions .....	13
Bouguer Gravity anomaly map .....	13
Aeromagnetic anomaly map .....	14
Regional Bouguer Gravity anomalies.....	14
Spectrometric data Analysis .....	16

Ternary Map.....	16
Geological Interpretation of Ternary map: .....	17
Major Observations:.....	17
<b>VISUALISATION AND ANALYSIS OF THE DATA .....</b>	<b>17</b>
Feature Importance .....	17
Integration Methodology: .....	18
1. Fuzzy Logic (Fuzzy Membership & Fuzzy Overlay) .....	18
2. Weighted Average (Weighted Overlay Tool).....	19
Workflow: Fuzzy Logic + Weighted Average in ArcGIS .....	19
<b>DEVELOPMENT OF AI/ML ALGORITHM.....</b>	<b>20</b>
Artificial Neural Network (ANN).....	22
<b>EVALUATION OF MODELS .....</b>	<b>23</b>
Manganese Model.....	23
ANN-1.....	23
Prospectivity Map .....	24
ANN-2.....	26
Gold Model .....	28
ANN-1.....	28
Prospectivity Map .....	28
ANN-2.....	30
Iron Model .....	31
Prospectivity Map .....	32
<b>UNSUPERVISED CLASSIFICATION OF THE FIELD DATA .....</b>	<b>33</b>
<b>DEPTH MODELING FOR MINERALISED BODIES.....</b>	<b>40</b>
1. Introduction.....	40
2. Data Preparation.....	41
2.1 Data used for the model preparation.....	41
2.2 Prism Grid Definition .....	43
3. Forward modeling Formulation .....	43
3.1 Initialization of Density Model .....	43
3.2 Prism Boundaries and prism centers.....	44
3.3 Sensitivity matrix calculation .....	44
<b>Integral representation of gravity anomaly.....</b>	44
Vectorization and Parallel computing.....	45
3.4 Gravity Prediction.....	46
4. Inversion Methodology .....	46
<b>4.1 Objective .....</b>	46

4.2 Conjugate Gradient Algorithm Overview.....	46
Steps of the Algorithm: .....	46
Results and Discussions.....	47
Radially averaged power spectrum of Gravity anomalies .....	47
Conclusion from the maps .....	48
<b>SOURCE CODE AND RESOURCES .....</b>	<b>49</b>
<b>REFERENCES.....</b>	<b>49</b>

## INTRODUCTION

Iron, copper, gold, manganese, nickel, Galena and other commercially viable commodities are among the metallic resources that are in high demand in India (GSI, SOP, 2011). The nation's industrialization and development depend on these metals. Kesler (2007) conducted a comparison between the global population and steel usage starting in 1700. The findings indicate a substantial increase in the global population since the start of the twenty-first century and the mining exploration sector has not been able to find new ore reserves at the same pace as global consumption in past decades (Goodway, 2012). There are two major issues that are emerging. First, the majority of base metal deposits close to the surface have already been found. Future research must look for deeper and hidden mineral deposits (Schmitt et al., 2003). Secondly, new and more effective exploration methods must be created. Drilling is a vital instrument in mineral discovery, while being the costliest. The necessity to protect our environment and promote sustainable development has sped up the development of novel geophysical technology for economical and non-destructive mineral prospecting in covered areas (e.g., Singh et al., 2019). This project focuses on developing machine learning-based prospectivity mapping algorithms for iron, copper, gold and other commodities in the predefined 39000 SKm area of Andhra Pradesh and Karnataka. It utilizes litho-geochemical, lithology, gravity and magnetic information and exploration reports provided by the Geological Survey of India (GSI).

### Regional Geology

Regional geology of the study area was prepared using ArcGIS 10.8.2 and has been shown in the figure below.

### Geology of Karnataka

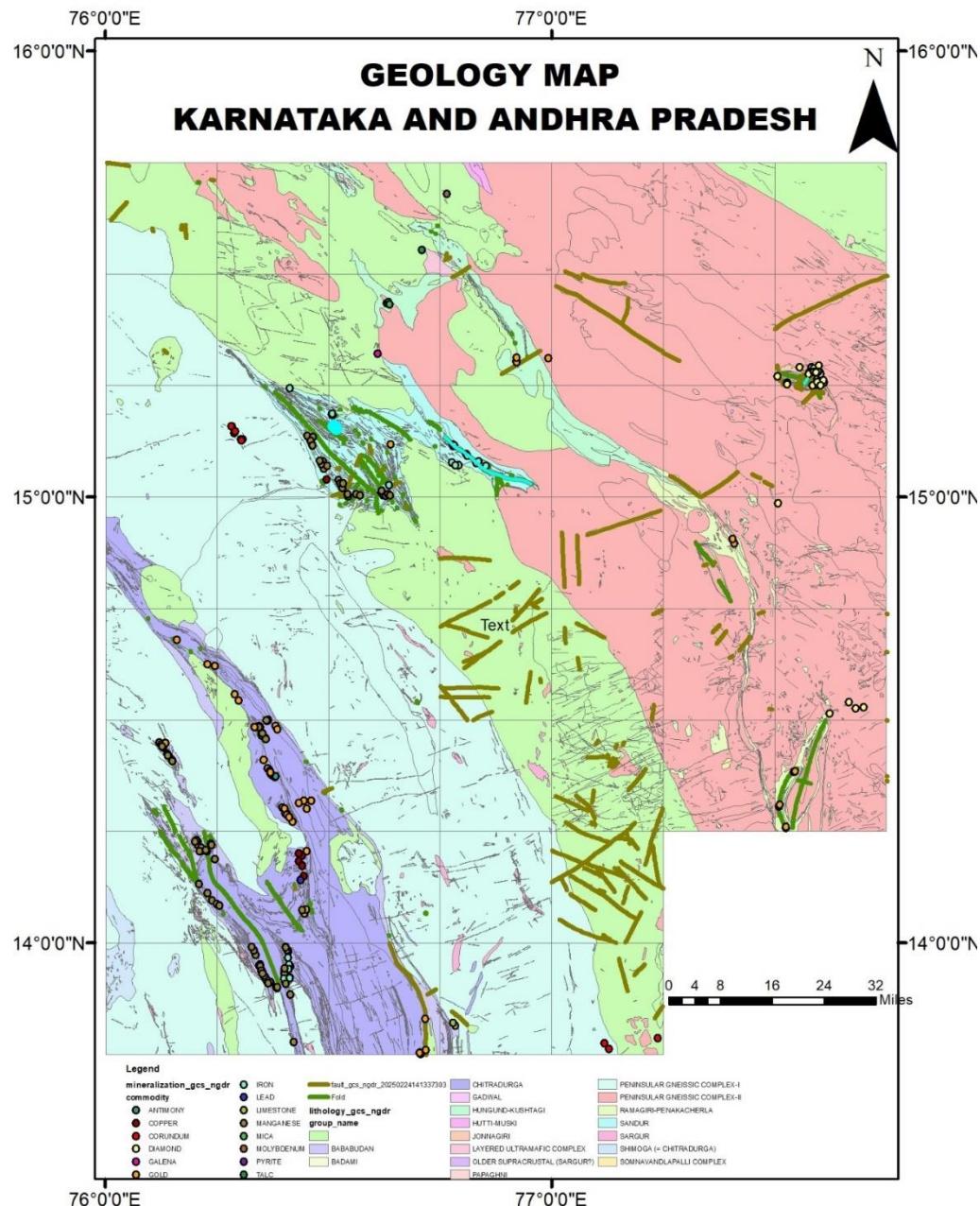
Karnataka can be broadly divided into three geological regions:

1. The Archean Complex (Older than 2.5 billion years) : a. composed of gneisses, schists and granites forms the foundation of Karnataka Geology. b. Found mainly in the Dharwar, Bellary and Kolar regions c. rich in minerals like gold, iron ore and manganese.
2. The Proterozoic formations (1.6-2.5 billion years old): a. Includes the Dharwar supergroup of rocks, which host significant mineral deposits b. Found in Chitradurga, Sandur and Bababudan regions. C. Contains important banded iron formation (BIFs).
3. The Deccan Traps (Cretaceous- Paleogene, ~ 66 million years old): consists of basaltic lava flows b. found in the northern districts like Bidar, Belgum and Gulbarga Forms a plateau with fertile black soil, suitable for cotton cultivation.
4. Sedimentary formation (Recent to Mesozoic) : Laterites and alluvium deposits are found in coastal and riverine areas. Form due to weathering of older rocks. Common in the Western ghats and along the Krishna and Kaveri river basins

### Geology of Andhra Pradesh

Andhra Pradesh consists of three major geological regions

1. Eastern Ghat mobile Belt (EGMB) : These rocks are of Archean to Proterozoic age. Located in the eastern part of the state. Composed of primarily of high grade metamorphic rocks such as khondalites, charnockites and granitic gneisses.
2. Cuddapah Basin: A prominent Proterozoic sedimentary basin covering parts of Kadapa, Kurnool and Anantapur districts. Contains rocks like quartzites, shales, dolomites and limestones.
3. Dharwar Craton: Composed of ancient Archean rocks such as granites and greenstone belts. Hosts gold, copper and other metallic ores.



Geology Map of the study area

## OBJECTIVES

The Objectives are defined by the problem statement of the competition **Hackathon 2025 on “Mineral Targeting using Artificial Intelligence”.**

1. Identification of new potential areas for exploration of critical minerals like REE, Ni-PGE, and copper, as well as other commodities like diamond, iron, manganese, and gold within a predefined 39,000 sq. km area in the states of Karnataka and Andhra Pradesh, India.
2. Emphasis on locating concealed & deep-seated mineralised bodies with depth modelling.
3. Developing AI/ ML algorithms for data cleaning, integration, modeling, and validation.
4. Generation of mineral predictive maps showing exploration targets visualised through maps, sections, etc.”

## NAME AND DETAILS OF PARTICIPANT AND COMPANY

### Team Members:

1. DR. SAUDAMINI SAHOO
2. NIHARIKA DAS
3. PIU DHIBAR

### Organisation:

Geological Survey of India

### Mail id

[saudamini.sahoo3@gmail.com](mailto:saudamini.sahoo3@gmail.com)

saudamini.sahoo@gsi.gov.in

## HARDWARE AND SOFTWARE USED

Google Collab GPU enabled version was used to write and run the program. Data was kept in google drive and was accessed from it through google drive. All the codes are done using the python language. For better clarity of Gravity Maps plotting Geosoft Oasis montaj 2024.2 data has been used. For preparation of different datasheets for Machine learning algorithm ArcGIS 10.8.2 was used.

## METHODOLOGY:

### 1. Data Collection:

Collected rock sample and litho-geochemical data from available exploration datasheets are systematically compiled and preprocessed. The corresponding predictor maps, incorporating a comprehensive range of geological, geochemical, and geophysical features to support subsequent analysis and modeling efforts.

## 2. Data Preparation:

The input data file is being prepared using the Arc-GIS software and was cleaned using python. The data preparation includes the following steps:

**(i) Lithology data:** The lithology data derived from 1:50,000 scale geological mapping is currently being sorted and cleaned using ArcGIS 10.8.2. This process involves organizing the dataset, correcting attribute inconsistencies, removing redundant or erroneous entries, and ensuring topological accuracy to prepare the data for subsequent spatial analysis and interpretation.

**(ii) Mineralogy data:** The available data is being intersected with the lithology dataset, specifically targeting lithological units known to host mineral deposits. This spatial intersection ensures that only the areas where the geological formations are associated with mineralization are retained for further analysis, thereby enhancing the relevance and accuracy of the results in the context of mineral exploration.

**(iii) Fault data:** The fault data, which contains geographic information on fault lines, is processed by applying a 1-kilometer buffer around each fault feature. This buffering creates a spatial zone of influence surrounding the faults. Subsequently, a spatial join is conducted between the buffered fault data and the lithology dataset, enabling the association of lithological attributes—such as rock types and geological units—with the areas in proximity to the faults. This integration facilitates a more comprehensive understanding of the geological context around fault zones.

**(iv) Fold data:** The fold data, which includes information on geological folds and their respective locations, is first subjected to a 1-kilometer buffer operation. This process creates a zone extending 1 km outward from each fold feature. Following this, a spatial join is performed between the buffered fold data and the lithology dataset. This allows for the association of lithological characteristics—such as rock types and formations—with the corresponding fold features within the specified buffer distance, enabling integrated geological analysis.

**(v) Geochemical data:** The geochemical data is also being spatially intersected with lithological units that correspond to known mineralized zones. This integration enables the identification of geochemical anomalies within specific lithologies that are favorable for mineral deposition, thereby enhancing the targeting accuracy for mineral exploration.

**(vi) Geophysical data:** The geophysical data, specifically ground gravity measurements, is spatially joined with lithological units known to host mineralization. Similarly, lithological units that do not exhibit evidence of mineralization are also spatially joined with the geophysical data. These two datasets—representing mineralized and non-mineralized lithologies—are then merged using the 'Union' tool under the 'Analysis' toolbox in ArcGIS. This process facilitates a comparative analysis of geophysical responses across different lithological and mineralization contexts.

The datasets generated from steps (i) to (vi) are being integrated using the 'Union' tool within the 'Analysis' toolbox in ArcGIS. This operation combines the spatial and attribute information

from all input layers into a single comprehensive dataset, enabling holistic spatial analysis and facilitating further interpretation across multiple geological and geospatial themes.

### 3. Data Cleaning

Further cleaning and processing of the input data (generated above) is done by utilizing Python libraries such as pandas and numpy, ensuring data consistency and quality, and applying appropriate normalization and transformation techniques to prepare the dataset for effective model training and analysis. Cleaning steps are there in the code.

## BRIEF DESCRIPTION OF THE AVAILABLE DATA

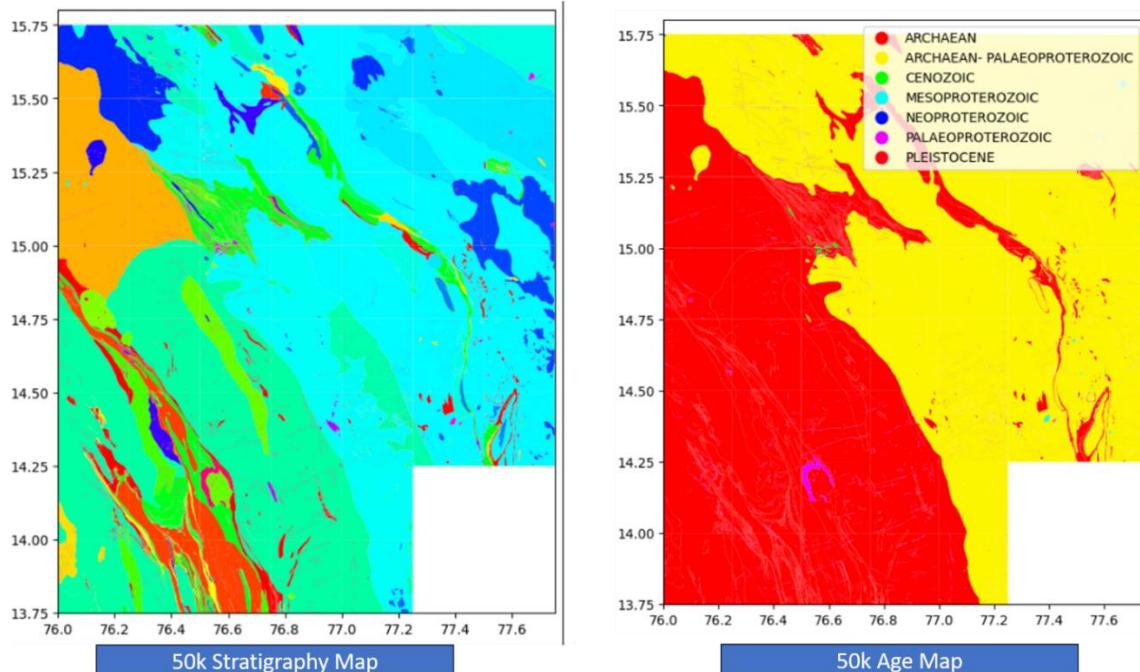
The brief description of the available data used are different geophysical, geological and geochemical data of Geological survey area in the pre-defined 39000 SKm area.

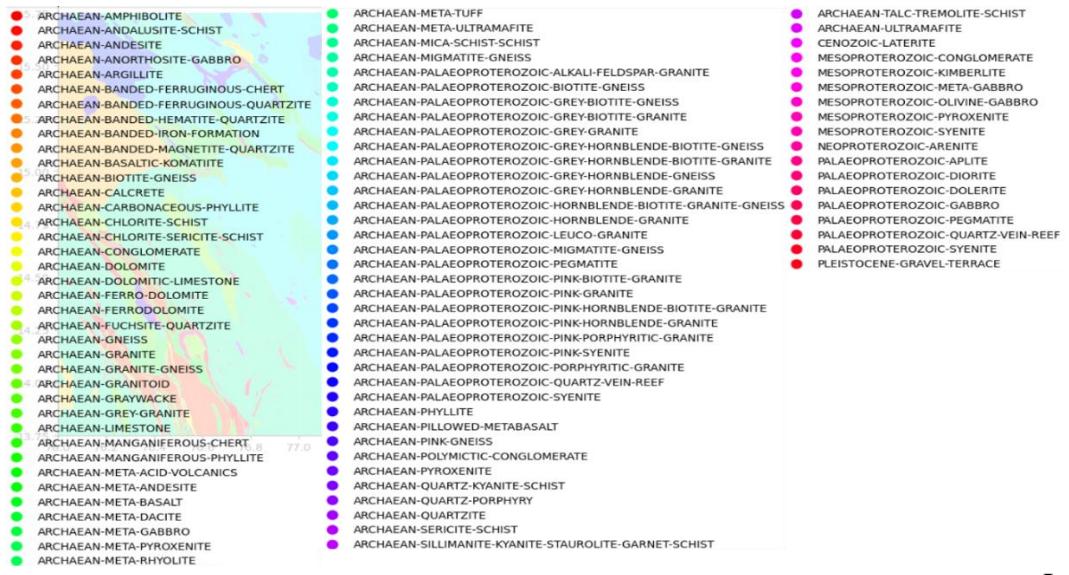
### 50K geological Data

The provided Geology map dataset contained various data folder- namely lithology (lithology\_gcs\_ngdr.shp), mineralization, mine\_quarry, faults etc. Each folder contained key geologic information and the data are analysed. For example, the lithology\_gcs\_ngdr shape file contained different key major attribute information like lithology, stratigraphy, supergroup, formation, age as an areal distribution in the form of a shape file. However all the attributes geometrics are kept as polygon in the shape files. An example of the geodataframe prepared from the lithology\_gcs\_ngdr shape file is described below:

#### 1. lithology\_gcs\_ngdr

The Figures show the distribution of geological age and stratigraphy

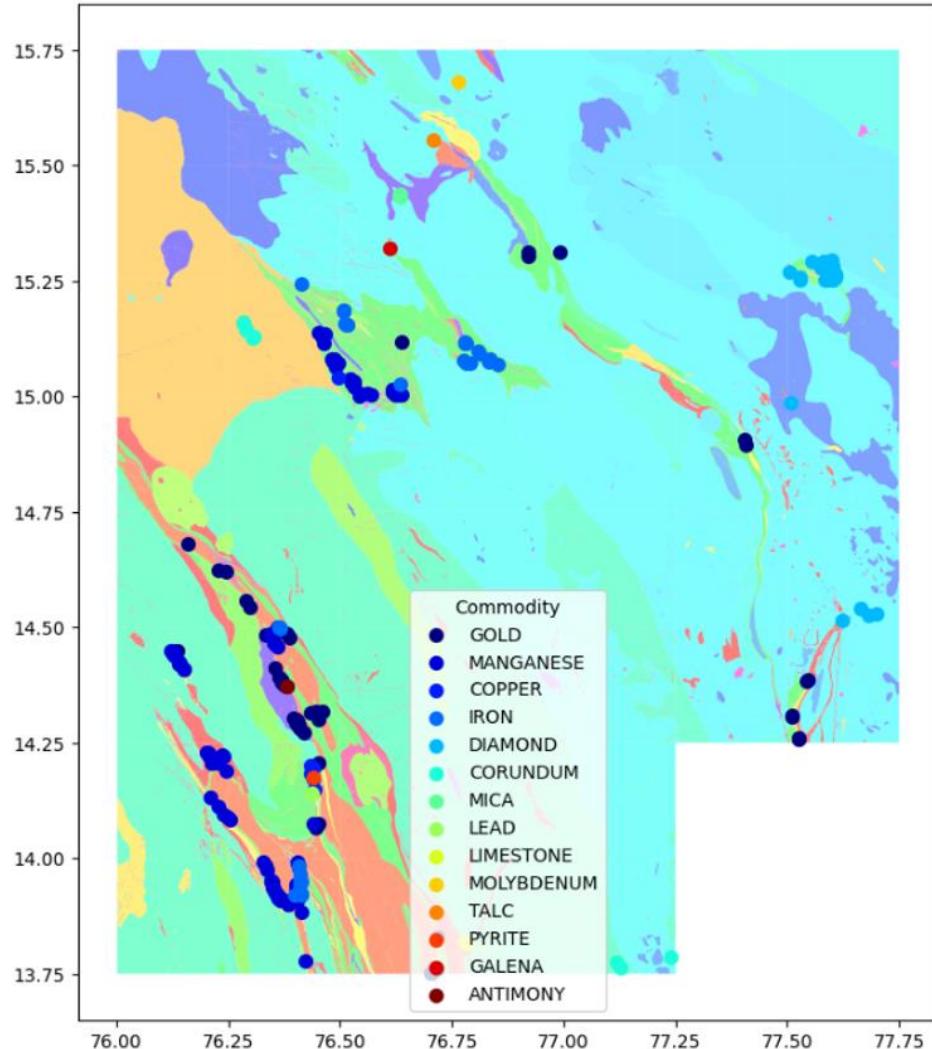




Legend

## 2. Commodity shape files

Commodity and Stratigraphy

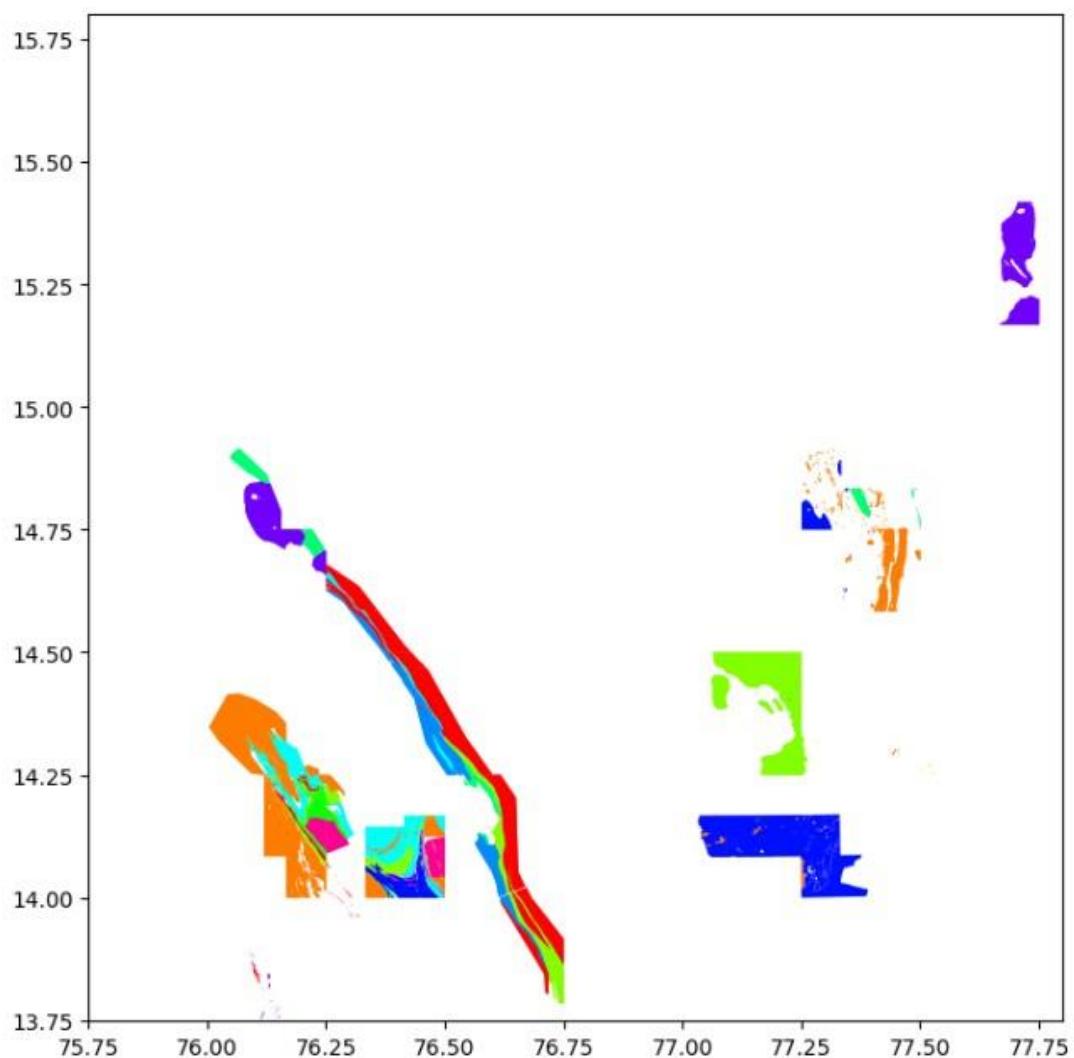


Using the ArcGIS 10.8.2. version both the lithology and commodity data points are combined. The intersection points between the commodity (Iron, Manganese, gold) are taken as points.

	x	y	lithologic	stratigr_1	supergroup	formation	group_name	commodity
0	76.7065	13.7515	META-BASALT	ARCHAEN-META-BASALT	DHARWAR	HIRIYUR	CHITRADURGA	GOLD
1	76.7035	13.7533	ARGILLITE	ARCHAEN-ARGILLITE	DHARWAR	HIRIYUR	CHITRADURGA	GOLD
2	76.7174	13.7606	BANDED IRON FORMATION	ARCHAEN-BANDED-IRON-FORMATION	DHARWAR	HIRIYUR	CHITRADURGA	GOLD
3	77.1281	13.7623	MIGMATITE GNEISS	ARCHAEN-MIGMATITE-GNEISS	PENINSULAR GNEISSIC COMPLEX	Nan	PENINSULAR GNEISSIC COMPLEX-I	CORUNDUM
4	77.1180	13.7750	MIGMATITE GNEISS	ARCHAEN-MIGMATITE-GNEISS	PENINSULAR GNEISSIC COMPLEX	Nan	PENINSULAR GNEISSIC COMPLEX-I	CORUNDUM

## 25K geological Data

25K geology data has been used for unsupervised Machine learning algorithm to find the different types of minerals.



## Geochemical Data Analysis and Interpretation:

This study presents a geochemical analysis of 10,004 stream sediment samples encompassing 73 elements, including major oxides, trace elements, and precious metals, supported by geospatial coordinates. The sediments are predominantly silicate-rich, as indicated by high SiO<sub>2</sub> content, with variable concentrations of Fe<sub>2</sub>O<sub>3</sub>, Al<sub>2</sub>O<sub>3</sub>, and carbonates. Trace elements such as Cr, Ni, Cu, and Zn show elevated levels in select samples, suggesting possible base metal mineralization. Gold concentrations reach up to 168 ppb, highlighting potential precious metal anomalies. These findings provide a foundation for identifying geochemical anomalies and prioritizing targets for further exploration.

A summary geochemical analysis of the dataset based on 10,004 stream sediment samples is as follows:

### 1. Major Oxides (%):

Oxide	Mean	Std Dev	Min	Max
SiO <sub>2</sub>	60.84	7.04	0.0	82.37
Al <sub>2</sub> O <sub>3</sub>	14.47	2.66	0.0	63.95
Fe <sub>2</sub> O <sub>3</sub>	5.56	4.61	0.0	81.12
CaO	3.14	2.52	0.0	21.30
MgO	1.90	1.24	0.0	12.52
Na <sub>2</sub> O	1.59	1.03	0.0	11.89
K <sub>2</sub> O	1.87	1.21	0.0	6.54
P <sub>2</sub> O <sub>5</sub>	0.11	0.07	0.0	1.61

- Observation:** The high mean silica (SiO<sub>2</sub>) suggests silicate-rich sediments, typical of felsic sources.

### 2. Trace Elements (ppm):

Element	Mean	Std Dev	Max
Ba	505	203	1864
Cr	98	71.7	1476
Zn	37	18.2	442
Cu	31.8	49.9	1908
Pb	17.2	11.3	483
Ni	44.1	30.3	528
As	3.06	6.40	150
Ag	20.0	37.0	677
Cd	35.1	58.0	500

- Observation:** Some samples have highly elevated levels of base metals (Cu, Zn, Ni), suggesting potential **mineralization zones**.

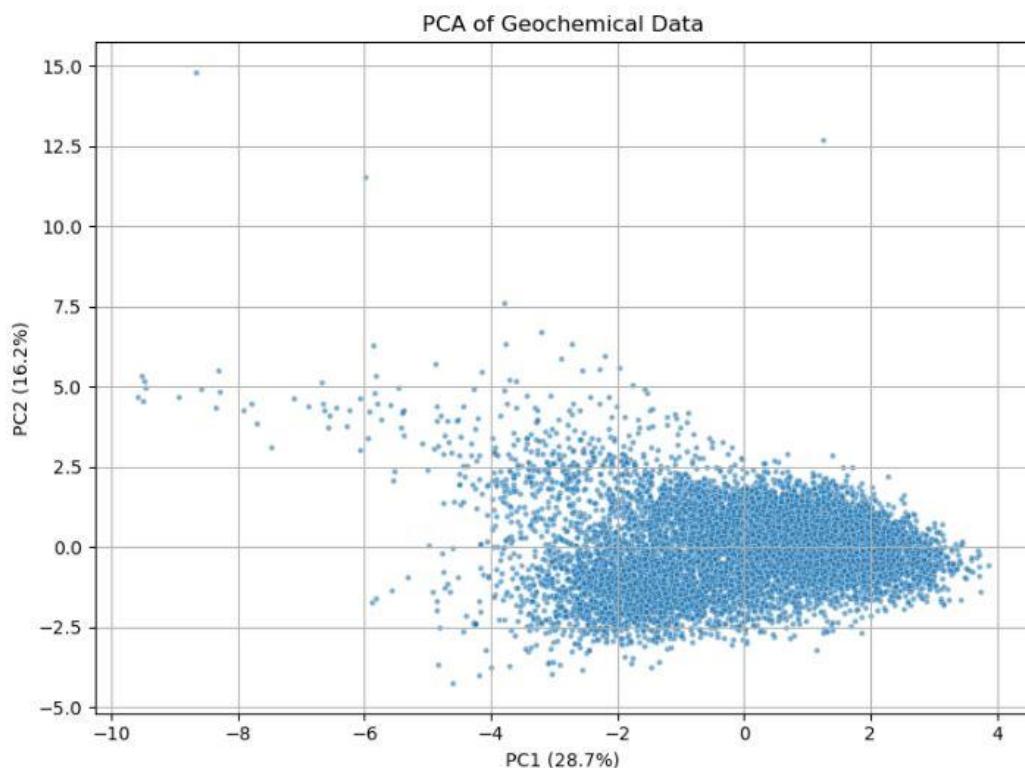
### 3. Precious Metals (ppb):

Element	Mean	Std Dev	Max
Au	1.59	4.39	168
Pt	0.10	0.31	8.0
Pd	0.09	0.29	9.1

- **Observation:** Gold (Au) shows significant variation, with a max of **168 ppb**, indicating possible gold anomalies worth further exploration

### Principal Component Analysis of Geochemical Data:

The plot is a PCA (Principal Component Analysis) scatter plot that visualizes the distribution of the geochemical stream sediment samples in terms of their first two principal components:



#### X-axis (PC1 – 28.7%):

Represents the first principal component, which explains 28.7% of the total variance in the dataset. This axis likely reflects the dominant geochemical process (e.g., lithological control or mineral composition like silicate vs. mafic content).

### **Y-axis (PC2 – 16.2%):**

Represents the second principal component, explaining an additional 16.2% of the variance. This component may reflect a secondary geochemical trend, such as hydrothermal alteration, weathering intensity, or anthropogenic input.

### **Data Points:**

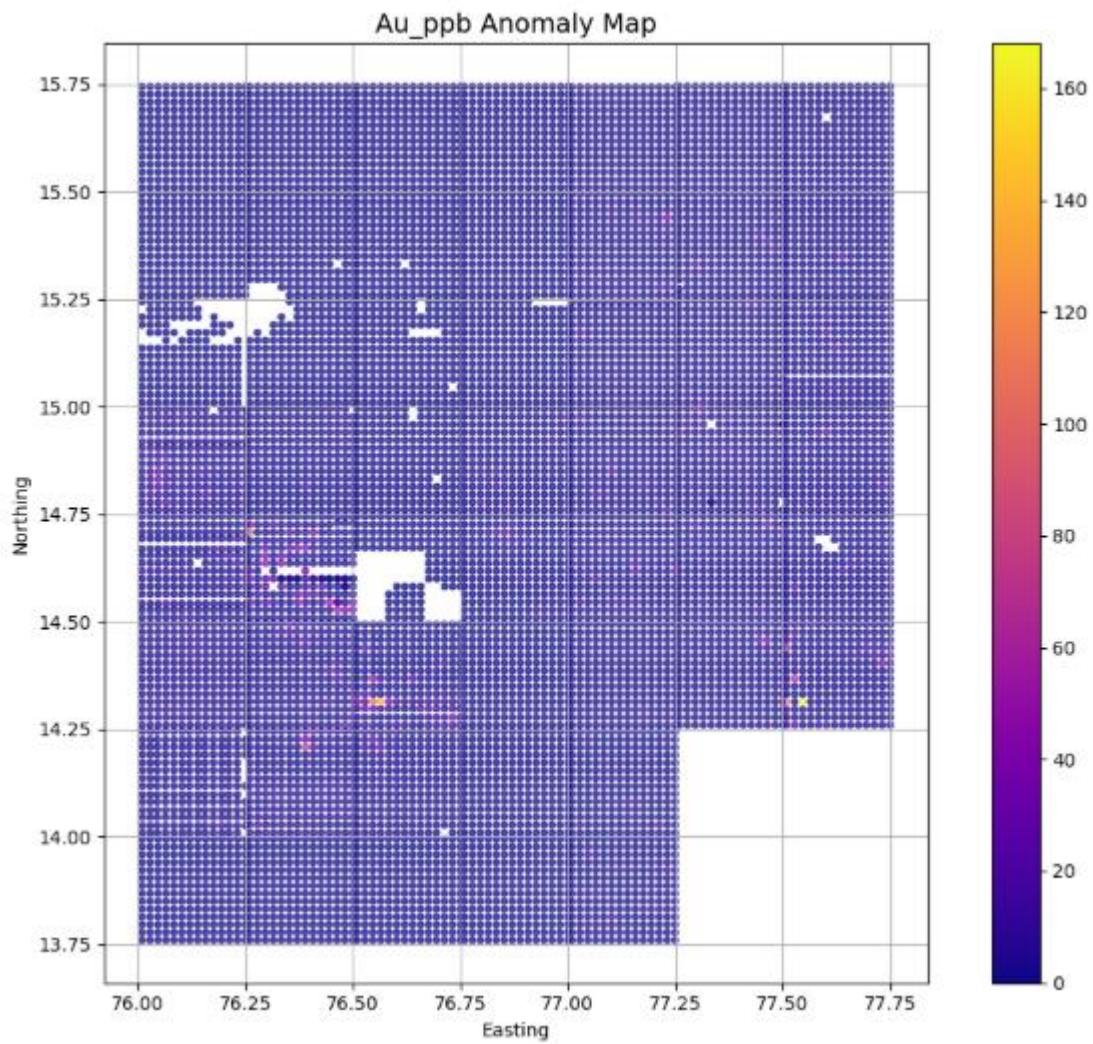
Each dot is a sediment sample projected into the PC1–PC2 space. The dense cluster on the right suggests a dominant compositional signature shared by most samples, while the more scattered points to the left and top may indicate anomalous or geochemically distinct samples.

### **Interpretation**

- Samples at the **extreme left (low PC1)** may be enriched in heavy elements or have different lithological sources (e.g., mafic-ultramafic terrains).
- Samples higher along **PC2** may reflect **elevated trace or precious metals**, weathering, or geochemical anomalies.
- The overall elliptical shape suggests a continuous variation rather than distinct groups, indicating **gradual changes in sediment chemistry** across the study area.

### **Anomaly Plot for Gold:**

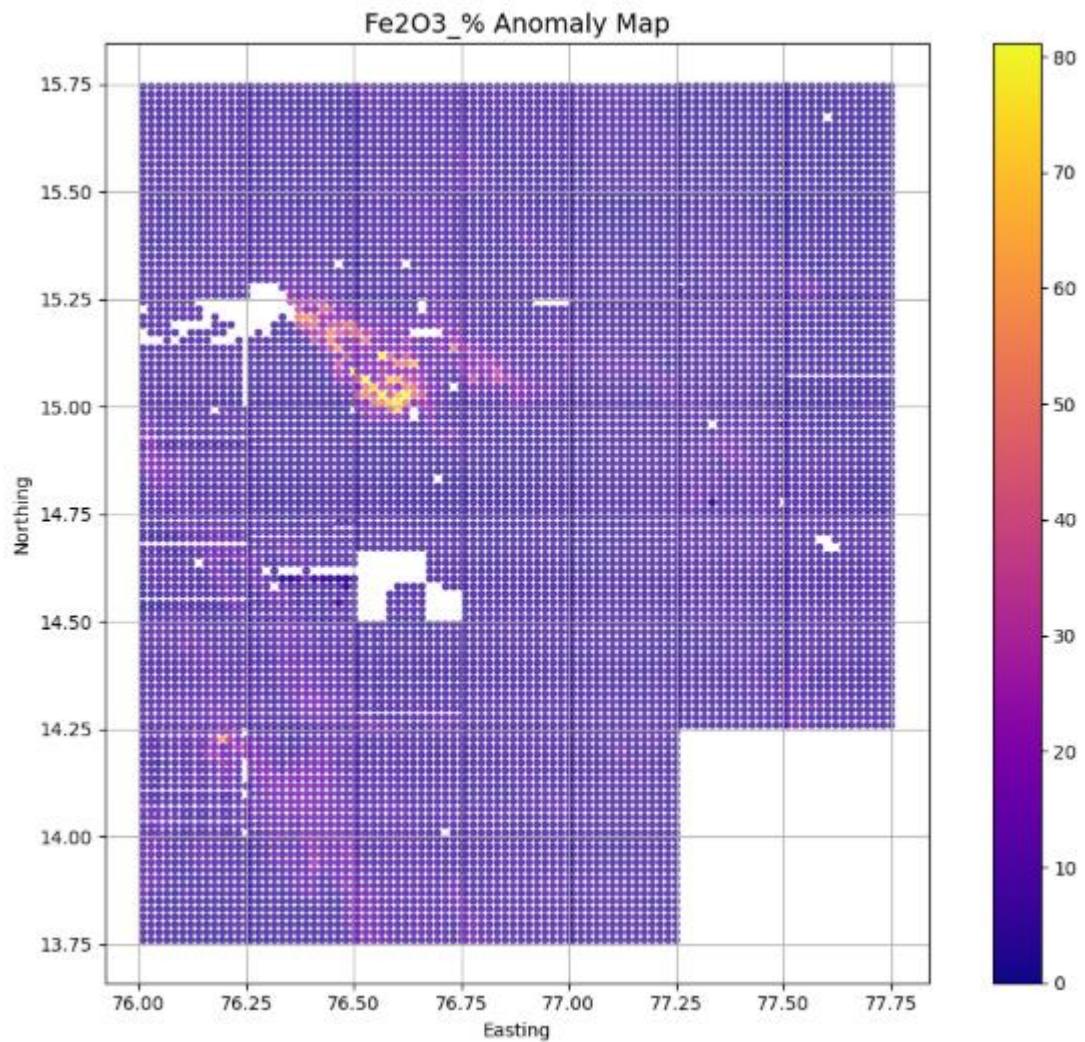
The region is predominantly dark blue or purple, meaning very low gold concentrations across most of the area. A few small bright spots (yellow to white) are scattered around — these represent localized high gold anomalies (Near Easting  $\approx$  76.5, Northing  $\approx$  14.5 and Near Easting  $\approx$  77.25, Northing  $\approx$  14.25, Smaller clusters in the northwest region). The isolated high-gold anomalies might indicate zones of interest for gold exploration (e.g., veins, deposits). The general low background suggests that mineralization is not widespread, but rather localized. These high-anomaly zones are likely **targets for further investigation**.



#### Anomaly Plot for Iron Oxide:

**Large High-Anomaly Zone:** Between approximately (Easting  $\approx$  76.25–76.75, Northing  $\approx$  15.0–15.25).

The Fe<sub>2</sub>O<sub>3</sub> anomaly is relatively strong and well-defined, suggesting the presence of iron-rich rocks or alteration zones. High Fe<sub>2</sub>O<sub>3</sub> concentrations often correspond to hydrothermal alteration, iron-rich lithologies (e.g., banded iron formations), or oxidized mineralized systems.



#### Comparison of Au and Fe<sub>2</sub>O<sub>3</sub> Map:

<b>Aspect</b>	<b>Observation</b>	<b>Meaning</b>
<b>High Au + High Fe<sub>2</sub>O<sub>3</sub> Overlap</b>	A few zones show stronger brightness and complexity (especially mid-right side).	These zones are <b>priority exploration targets</b> — possible <b>hydrothermal gold systems</b> related to iron alteration.
<b>High Au without Strong Fe<sub>2</sub>O<sub>3</sub></b>	Some Au spots are isolated without strong Fe <sub>2</sub> O <sub>3</sub> background.	Could indicate <b>gold mineralization independent of large-scale iron alteration</b> — possibly narrow veins.
<b>High Fe<sub>2</sub>O<sub>3</sub> without High Au</b>	Several areas have strong Fe <sub>2</sub> O <sub>3</sub> but no obvious Au anomalies.	May represent <b>barren iron oxides</b> (e.g., laterite, gossan) — needs field validation.
<b>Low Fe<sub>2</sub>O<sub>3</sub> + Low Au</b>	Dominant in large areas.	Probably <b>non-prospective ground</b> — lower priority.

## Analysis of Geophysical Data

### Introduction

Gravity and magnetic methods are very useful tools for delineating various litho-units having density and susceptibility contrasts. Various contour patterns reflect lithological disposition and structural features. In addition, quantitative analysis of gravity and magnetic data is also done to infer the depth of occurrence of major formations and structures.

### Geophysical Data processing and preparation of maps

The Oasis Montaj (version 2024.2) package of Geosoft Inc. has been used for the processing and generation of gravity and magnetic anomaly maps. The Bouguer gravity anomaly and aeromagnetic anomaly values have been projected through UTM projection system with WGS 84 datum in zones 43N and, a digital base map has been prepared. Bouguer gravity and aeromagnetic anomaly values have been re-gridded at 500 m intervals and, contours have been drawn on 2 mGal and 100 nT intervals, respectively to prepare anomaly maps. The gridded gravity and magnetic data have been filtered by digital filtering techniques using Fast Fourier Transform (FFT) (Bhattacharya, 1967) to enhance the response of a geological structure, suppressing the effect of noise. Processed maps such as residual gravity and magnetic anomaly maps, gravity and magnetic first-order derivative maps have been generated from the filtered data. Radially Averaged Power Spectra of Bouguer gravity anomaly have been obtained by transforming to the frequency domain, and thus, depth corresponding to different frequencies has been obtained from the power spectrum.

### Results and Discussions

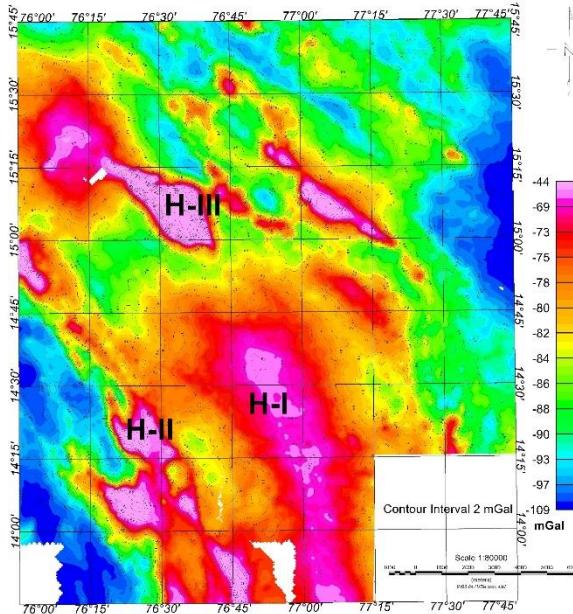
Qualitative and quantitative analyses of gravity and magnetic data have been done based on different anomaly maps, depth solutions, etc. to characterize the area with special attention to the subsurface features.

#### Bouguer Gravity anomaly map

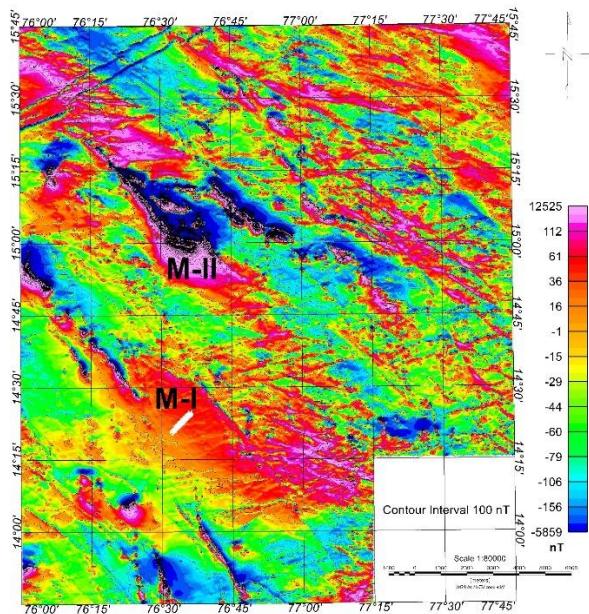
Bouguer gravity anomaly map brings out the density contrast amongst geological features and formations. It enables us to distinguish individual units from an assemblage based on their characteristic gravity signatures. Basic and ultrabasic rocks are denser than felsic and sedimentary rocks and hence yield higher gravity signals.

Bouguer gravity anomaly varies between -109 mGal and -44 mGal with a variation of the order of 65 mGal and is shown below. The prominent features of the Bouguer anomaly contour map are the elongated high zones (H-I and H-II) where H-I corresponds to the Ramagiri-Penakacherla schist belt and H-II corresponds to the Chitradurga schist belt. H-I is aligned to NW-SE direction in the south-central part, flanked by H-II to the NW-SE direction in the southwestern part. H-I and H-II correspond gravity anomalies of roughly 20 mGal relief with respect to the regional background and appear to have an approximate width of 35 km and 20 km respectively. The conspicuous changes in the orientation of H-II as compared to H-I indicate the change in the trend of the schist belt towards the west. Another distinct gravity high (H-III) with magnitude of the order of 40 mGal, width of about 20 km, trending in NW-SE

direction, has been picked up in the northwestern part. This gravity high corresponds to the Kushtagi schist belt. It is evident that the gravity high zones (H-I and H-II) are discontinuous. On the other hand, the lows on either side of H-I and H-II show up as Bouguer gravity anomalies of  $-90$  mGal to  $-109$  mGal magnitudes with an accompanying change in the shape of the contours. The gravity lows in the eastern and southwestern part may be attributed to low-density younger granites.



Bouguer Gravity anomaly map



Aeromagnetic anomaly map

#### Aeromagnetic anomaly map

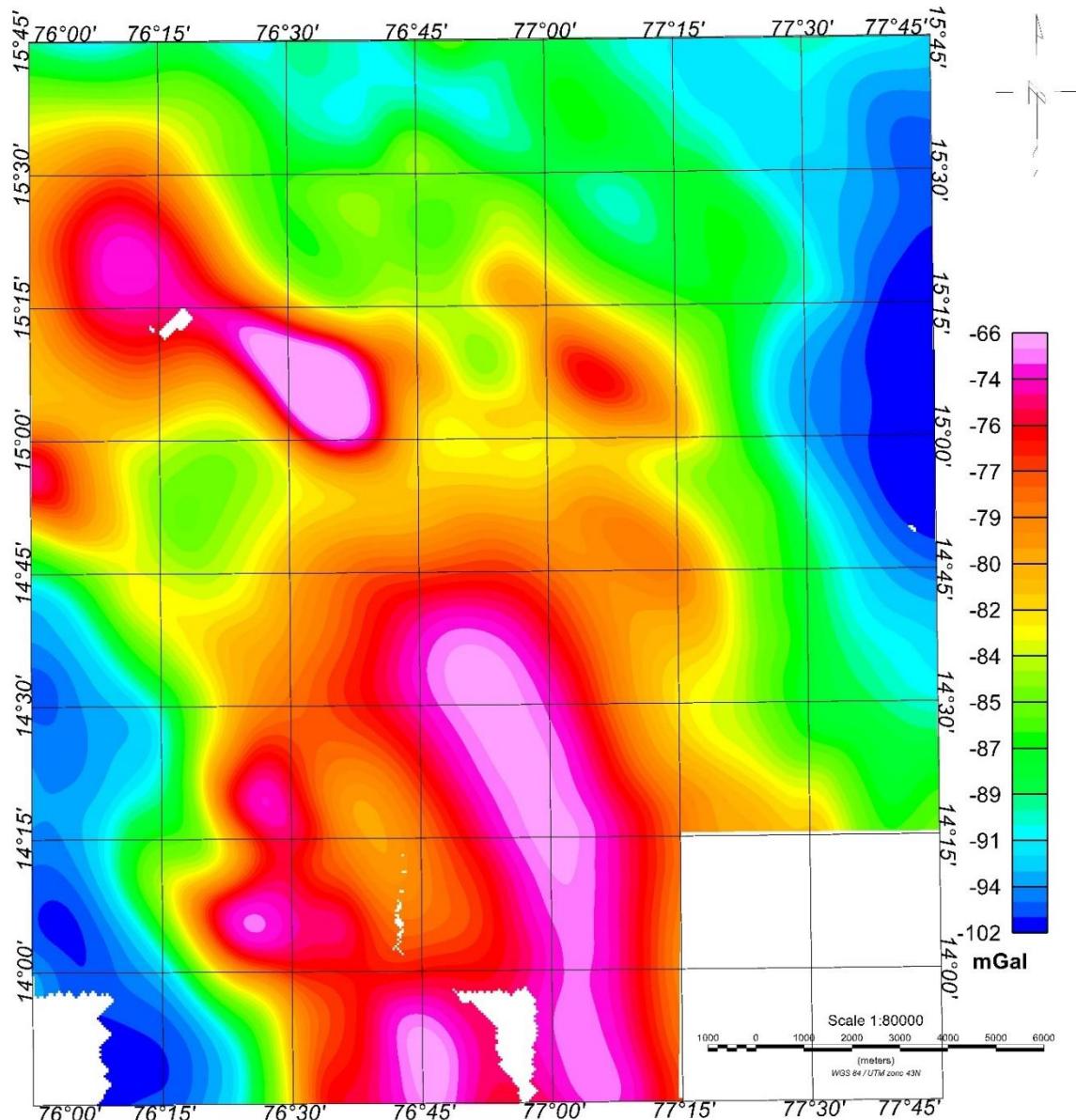
Aeromagnetic anomaly map has been prepared for contour intervals of  $100$  nT. The magnetic anomaly values of the study area range from about  $-5859$  nT to  $12525$  nT with a total variation of the order of  $18384$  nT and is shown above. Some distinct anomaly zones can be identified from the aeromagnetic map. The anomaly zone, M-I, at the southcentral part of the study area is a prominent high, which indicates the presence of magnetic minerals in the Chitradurga schist belt. A conspicuous bipolar magnetic anomaly, M-II, running in the NW–SE direction in the northwest part of the study area is observed over the Kushtagi schist belt. This signifies the presence of magnetic minerals in the Kushtagi schist belt. There are other high frequency anomalies with low and high amplitudes.

Both gravity and magnetic maps collaborate with each other and represents the mineralized zones nicely.

#### Regional Bouguer Gravity anomalies

For enhancement of the anomalies pertaining to regional features in suppression of the responses of local features, Upward continuation filter is used. It calculates the potential field at an elevation higher than at which the field is measured. Gravity field from an upward

continuation level of 'Z' retains anomalies from sources below a depth of  $Z_0 = (1/2) * Z$ , whereas sources above are more attenuated (Jacobsen, 1987; Lyngsie et al, 2006). The regional gravity anomaly has been accepted at cutoff wavelength 6 km where all the local anomalies have been removed. The regional gravity anomaly map has been represented in below figure. The regional trend of the anomalies is in the NW-SE direction. The regional gravity high at the southcentral part over the Ramagiri-Penakacherla schist belt and over the Kushtagi schist belt in the northwestern part is easily traceable up to 6 km upward continued map, which suggests the continuation of the anomaly sources up to a depth of 3 km or more.

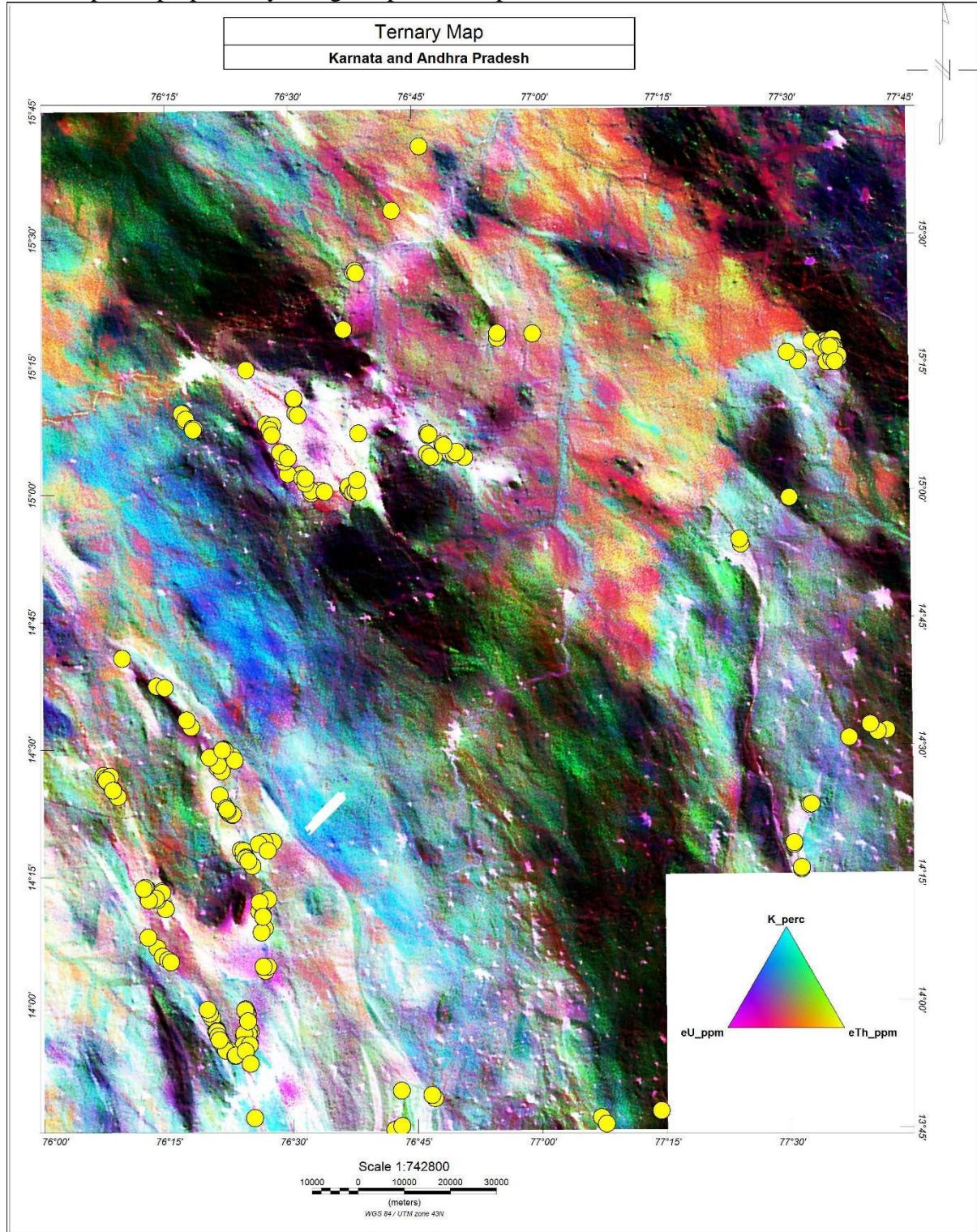


Regional Bouguer Gravity anomaly map

## Spectrometric data Analysis

### Ternary Map

This map was prepared by using the provided spectrometric data.



This is a **Ternary Radiometric Map** covering parts of **Karnataka and Andhra Pradesh**. It uses **radiometric (gamma-ray spectrometry) data** — mapping natural radiation from rocks and soils. The **ternary color scheme** represents:

- **K (potassium) content** – at one vertex of the triangle.
- **eU (equivalent uranium)** – at another vertex.
- **eTh (equivalent thorium)** – at the third vertex.

#### Geological Interpretation of Ternary map:

Feature	Interpretation
Pink/White Zones	Possible hydrothermal alteration zones (e.g., potassium metasomatism), often linked to <b>mineralization</b> (gold, base metals).
Green Zones	Potential <b>uranium-rich</b> areas — may coincide with fault structures.
Blue Zones	Thorium-rich — possibly heavy mineral concentrations or certain granitoids.
Sample Clusters	Exploration teams have <b>focused sampling</b> around <b>altered or anomalous areas</b> , suggesting prior geological interest.

The dense clustering of sample points in potassium-altered (pink/white) zones suggests active exploration, possibly targeting hydrothermal systems favorable for gold and uranium mineralization.

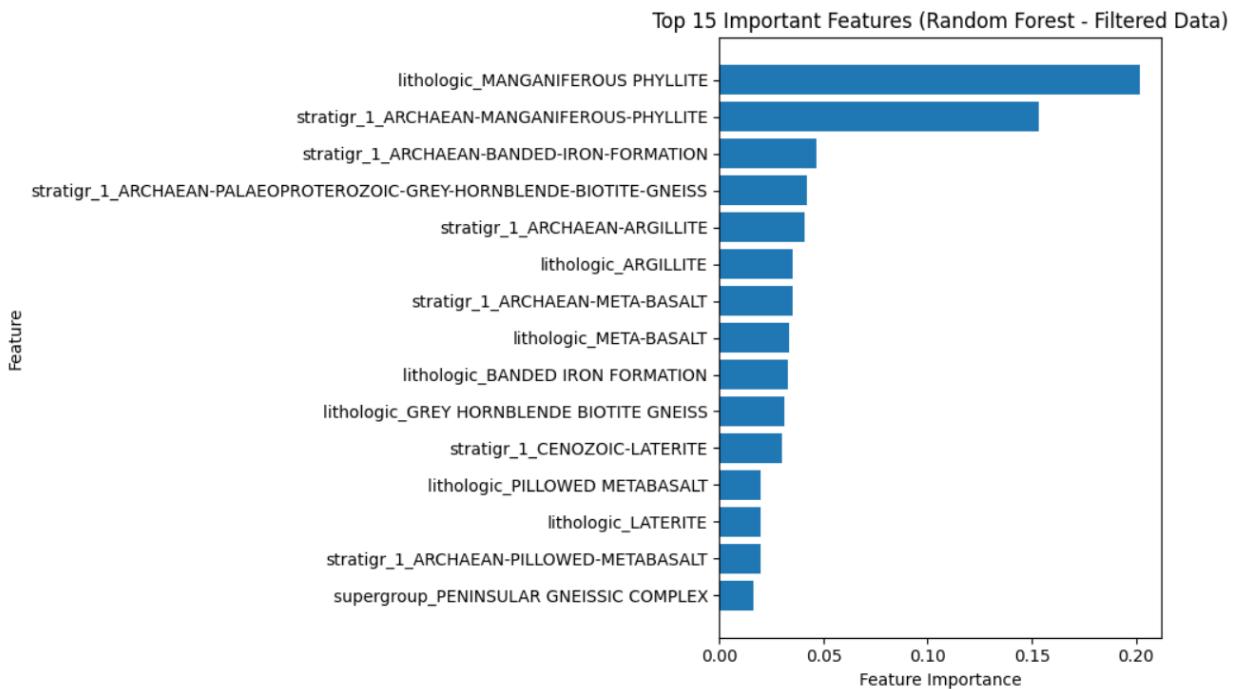
#### Major Observations:

Observation	Interpretation
Gold and Iron mineral cluster over pinkish/white zones in the ternary map.	These zones are potassium-enriched (K%), indicating hydrothermal alteration — a common host for gold mineralization.
Iron enrichment is also stronger where K-alteration is observed.	Iron oxides often form in oxidized alteration systems related to gold (especially IOCG-type or epithermal systems).
Anomalies align along certain linear features (northwest-southeast direction).	Likely faults or shear zones, critical pathways for hydrothermal fluids transporting gold and other minerals.
Low uranium/thorium zones (darker in background) correlate poorly with anomalies.	Suggests that radioelement depletion is not a strong indicator here — potassium is the key guide.

## VISUALISATION AND ANALYSIS OF THE DATA

#### Feature Importance

The Random Forest algorithm was used on 50K-geological data to find the top 15 important features to identify the commodity data points.



### Integration Methodology:

Data integration using fuzzy logic and weighted average is commonly done in ArcGIS 10.8.2. in spatial multi-criteria decision analysis (MCDA) — especially for suitability analysis, risk mapping, environmental modeling, or resource evaluation. In ArcGIS, this is achieved using tools like the Fuzzy Membership, Fuzzy Overlay, and Weighted Overlay tools available in the Spatial Analyst extension.

Two Main Tools in ArcGIS for this Integration are:

#### 1. Fuzzy Logic (Fuzzy Membership & Fuzzy Overlay)

Fuzzy logic is used to handle imprecise or uncertain criteria, converting input raster data into values between 0 and 1 (degrees of membership in a fuzzy set).

- **Fuzzy Membership Tool:** Converts input raster data (e.g., slope, elevation) into a fuzzy membership map.
  - Functions: *Linear, Gaussian, Small, Large*, etc.
  - Output: Raster where each cell has a value between 0 (non-membership) and 1 (full membership).
- **Fuzzy Overlay Tool:** Combines multiple fuzzy membership rasters using fuzzy operators like:
  - **Fuzzy AND** (minimum)
  - **Fuzzy OR** (maximum)
  - **Fuzzy GAMMA** (controlled blend between AND/OR)
  - **Fuzzy SUM** (compensatory increase)

## 2. Weighted Average (Weighted Overlay Tool)

The **Weighted Overlay Tool** combines multiple input rasters that have been reclassified to a common scale (e.g., 1–9), using user-defined weights for each layer based on importance or confidence.

Steps:

- Reclassify each raster to a common scale (using *Reclassify* tool).
- Assign weights (percentages summing to 100%).
- Combine using the Weighted Overlay tool.

### Workflow: Fuzzy Logic + Weighted Average in ArcGIS

1. **Preparation of Raster Layers** for lithology , mineral, fold, fault, Bouguer gravity data, Residual Bouguer gravity data, Ternary data (Spectrometric data).

2. **Apply Fuzzy Membership**

Used **Fuzzy Membership** tool to convert each raster into a fuzzy scale (0–1), based on how suitable or relevant it is for the target application.

3. **Use Fuzzy Overlay**

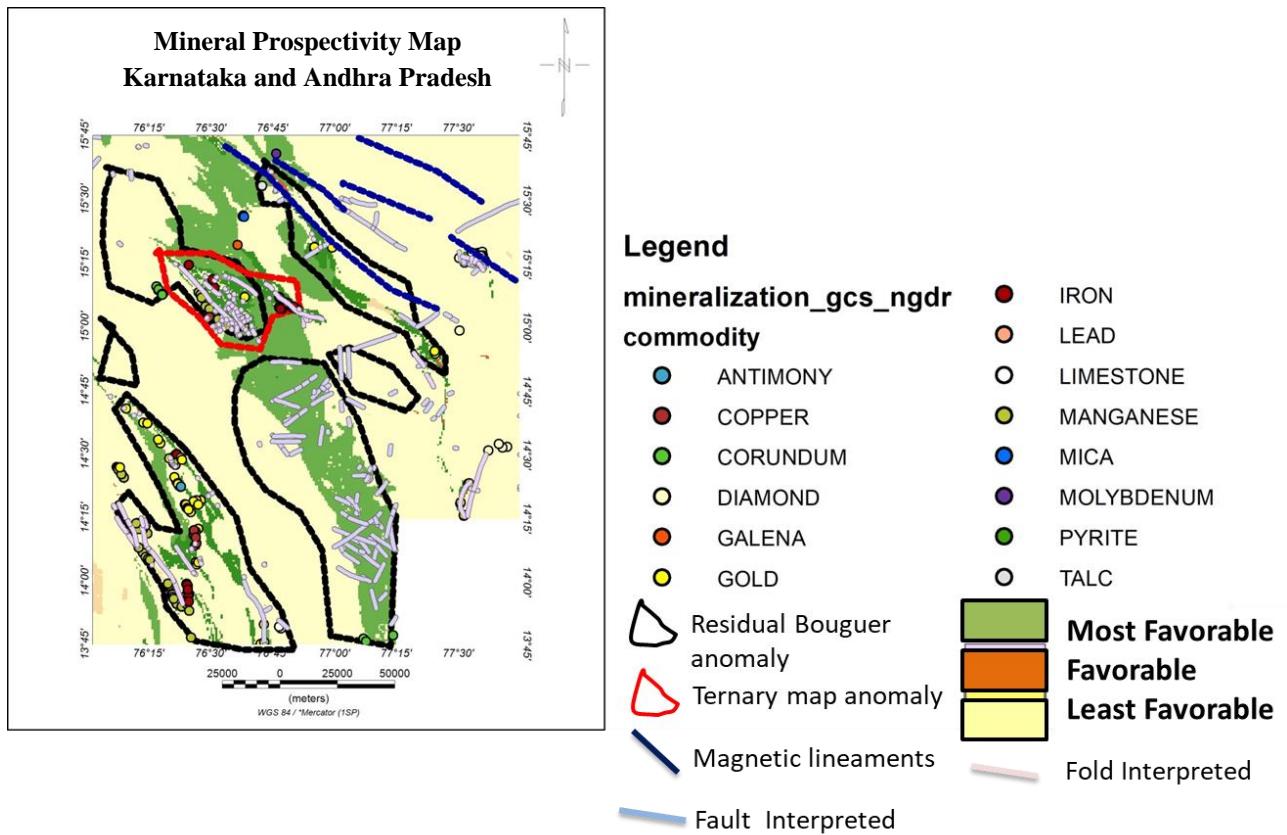
Combine these fuzzy layers using the **Fuzzy Overlay** tool to handle uncertain or subjective criteria using fuzzy logic with Gamma=0.9.

4. **Weight:** Assign importance and produce factor map. The following weightage are given:

- (i) Geological Factor Map- Lithology (9)
- (ii) Geophysical Factor Map- Bouguer gravity anomaly (8)
  - Residual Bouguer anomaly map (8)
  - Ternary map (7)
  - Aeromagnetic map (7)
- (iii) Structural Factor Map- Fault (7)
  - Fold (7)
- (iv) Geochemical Factor Map- (Fe<sub>2</sub>O<sub>3</sub>, MgO, Cr, Cu, Ag)

5. **Apply Weighted Overlay**

Different data sources have varying importance. By applying the **Weighted Overlay Tool** using reclassified rasters and user-defined weights, the output will be a mineral prospectivity map as shown in figure below.



## DEVELOPMENT OF AI/ML ALGORITHM

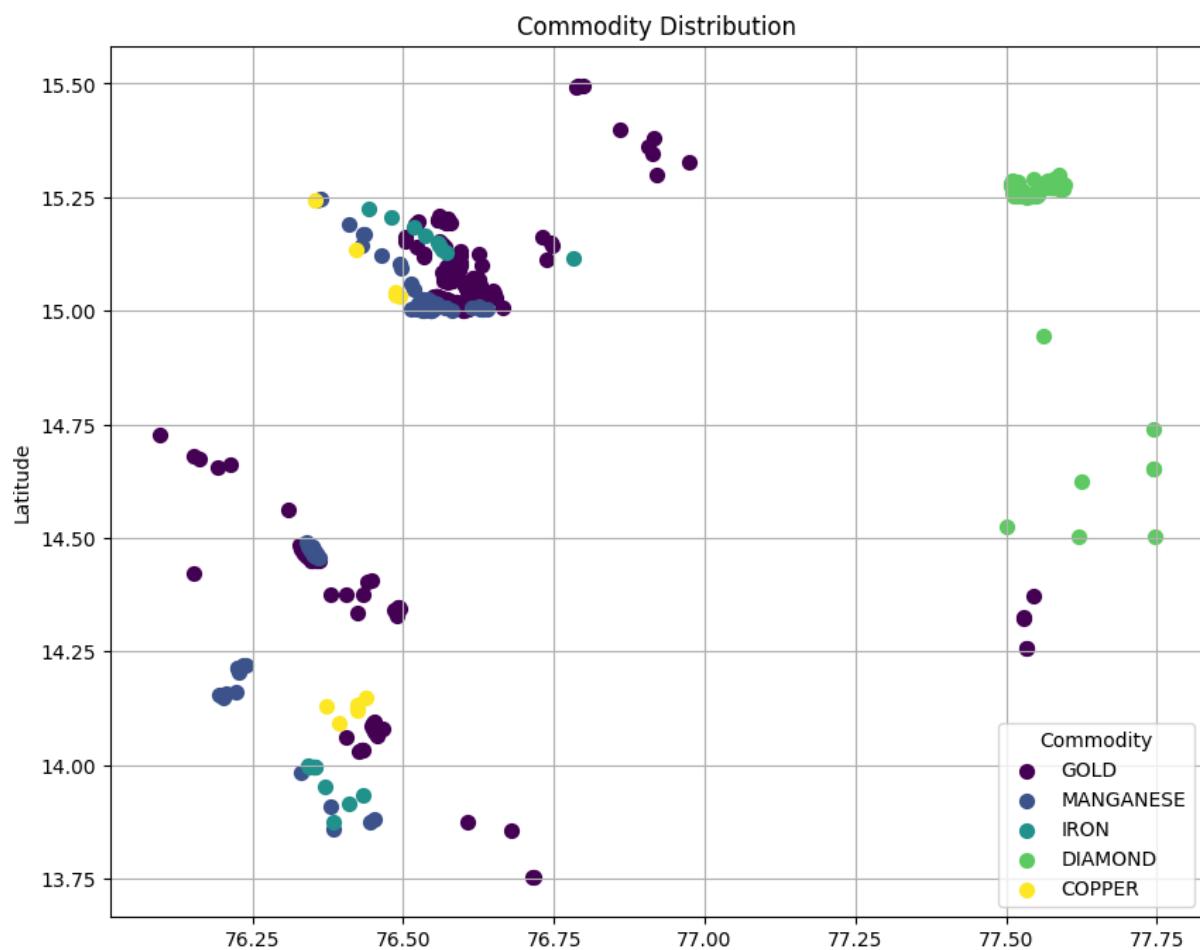
Separate model building approaches have been executed for prospectivity mapping, using training data where commodity points have been found in the shared geology shape files to create machine learning models for manganese, gold, and iron.

The Artificial Neural Network (ANN) algorithms were employed in the investigation. The data's important features have been found using the decision tree model.

These algorithms were selected because of their shown ability to handle complex, high-dimensional data and their ability to produce accurate, trustworthy forecasts.

Below, we explain how each method works and why it is crucial to our prospectivity modelling endeavours.

COMMODITY	POINTS
<b>GOLD</b>	301
<b>DIAMOND</b>	203
<b>MANGANESE</b>	127
<b>IRON</b>	33
<b>CORUNDUM</b>	13
<b>COPPER</b>	11
<b>TALC</b>	3
<b>MOLYBDENUM</b>	2
<b>MICA</b>	1
<b>GALENA</b>	1
<b>LIMESTONE</b>	1



## Artificial Neural Network (ANN)

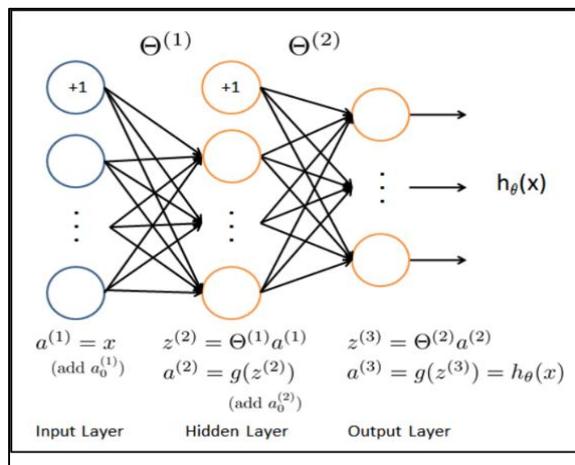
An Artificial Neural Network (ANN) model has been developed to classify and predict the presence of specific mineral commodities—namely Manganese, Gold and Iron. For each individual commodity model, a binary classification approach is adopted. For example, in the Manganese model, only data points explicitly associated with Manganese occurrences are

treated as positive examples, while all other data points corresponding to other commodities (such as Gold and Iron) are labeled as negative examples. This method of preparing the target variable ensures that the ANN focuses solely on identifying patterns relevant to the presence of the specific commodity under study. A similar strategy is followed for creating dedicated models for Iron, and Copper, allowing each model to specialize in recognizing its respective target commodity. The dataset used for training these models is split into an 80-20 ratio, with 80% used for training and the remaining 20% reserved for testing and validation.

The architecture of the ANN model used for this classification task comprises two hidden layers, culminating in an output layer with a sigmoid activation function. This design enables the network to output probabilities indicating the likelihood of the presence of a particular commodity at a given data point. The modeling process is carried out in two stages to evaluate the impact of different input features. In the first stage, only geochemical data and geophysical data—specifically Bouguer gravity anomaly values—are used as inputs (**ANN-1**). In the second stage, lithological information and geological age data are included alongside the geochemical and geophysical inputs (**ANN-2**). This two-step approach helps to assess whether the inclusion of geological context improves the predictive performance of the ANN. The model is designed to learn complex relationships among the various input attributes and produce probabilistic predictions that can support mineral exploration efforts. Predictive maps for both the models are included in the report.

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 128)	3,456
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 64)	8,256
dropout_1 (Dropout)	(None, 64)	0
dense_2 (Dense)	(None, 1)	65

ANN- 1: Total input feature 27



ANN Structure modified after  
Andrew Ng

The architecture of the ANN model used for this

classification task comprises two hidden layers, culminating in an output layer with a sigmoid

activation function. This design enables the network to output probabilities indicating the

likelihood of the presence of a particular commodity at a given data point. The modeling

process is carried out in two stages to evaluate the impact of different input features. In the first

stage, only geochemical data and geophysical data—specifically Bouguer gravity anomaly

values—are used as inputs (**ANN-1**). In the second stage, lithological information and

geological age data are included alongside the geochemical and geophysical inputs (**ANN-2**).

This two-step approach helps to assess whether the inclusion of geological context improves

the predictive performance of the ANN. The model is designed to learn complex relationships

among the various input attributes and produce probabilistic predictions that can support

mineral exploration efforts. Predictive maps for both the models are included in the report.

Layer (type)	Output Shape	Param #
dense_18 (Dense)	(None, 64)	4,928
dense_19 (Dense)	(None, 32)	2,080
dense_20 (Dense)	(None, 2)	66

ANN- 2: Total input feature 77

## EVALUATION OF MODELS

### Manganese Model

#### ANN-1

The model is run thrice. The accuracy and error score mentioned below are the average scores obtained from three different sets of training and testing data. The final results are mentioned in the code shared **in the github**:

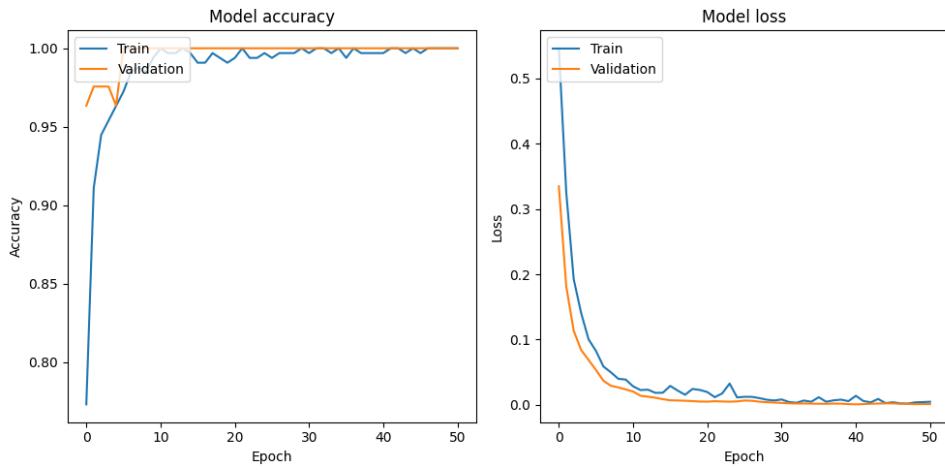
Training Accuracy	0.9909
Validation Accuracy	0.9902
Training Error	0.0024
Validation Error	0.0013

An example of the initial results is described in the figure below. The total number of iteration was set to 100 with an **EarlyStopping callback** in **TensorFlow (specifically Keras)**. The **callback** stops training a neural network **early** when a monitored metric (validation loss, accuracy) stops improving. The network shows saturation at iteration number showing the optimum use of network.

```
Epoch 1/100
11/11 ━━━━━━━━━━ 5s 78ms/step - accuracy: 0.6673 - loss: 0.6140 - val_accuracy: 0.9634 - val_loss: 0.3346
Epoch 2/100
11/11 ━━━━━━━━ 0s 21ms/step - accuracy: 0.9103 - loss: 0.3265 - val_accuracy: 0.9756 - val_loss: 0.1811
Epoch 3/100
11/11 ━━━━━━ 0s 21ms/step - accuracy: 0.9522 - loss: 0.1921 - val_accuracy: 0.9756 - val_loss: 0.1136
Epoch 4/100
11/11 ━━━━ 0s 18ms/step - accuracy: 0.9573 - loss: 0.1427 - val_accuracy: 0.9756 - val_loss: 0.0840
Epoch 5/100
11/11 ━━━ 0s 25ms/step - accuracy: 0.9632 - loss: 0.1040 - val_accuracy: 0.9634 - val_loss: 0.0689
```

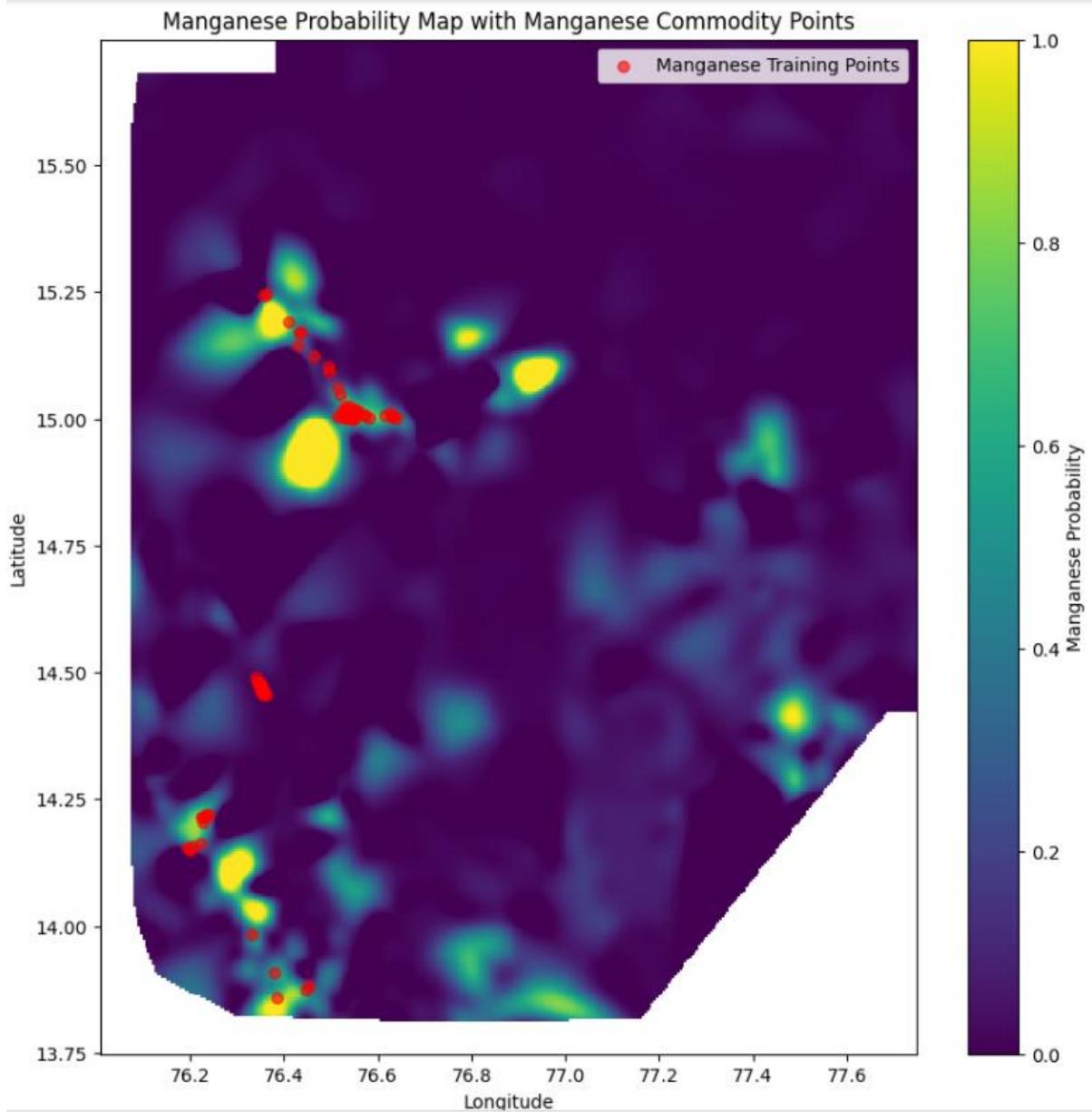
```
# Build Keras Sequential ANN model
model = Sequential([
    Dense(128, activation='relu', input_shape=(X_train.shape[1],)),
    Dropout(0.3),
    Dense(64, activation='relu'),
    Dropout(0.3),
    Dense(1, activation='sigmoid') # Binary classification
])

model.compile(optimizer=Adam(learning_rate=0.001), loss='binary_crossentropy', metrics=['accuracy'])
```

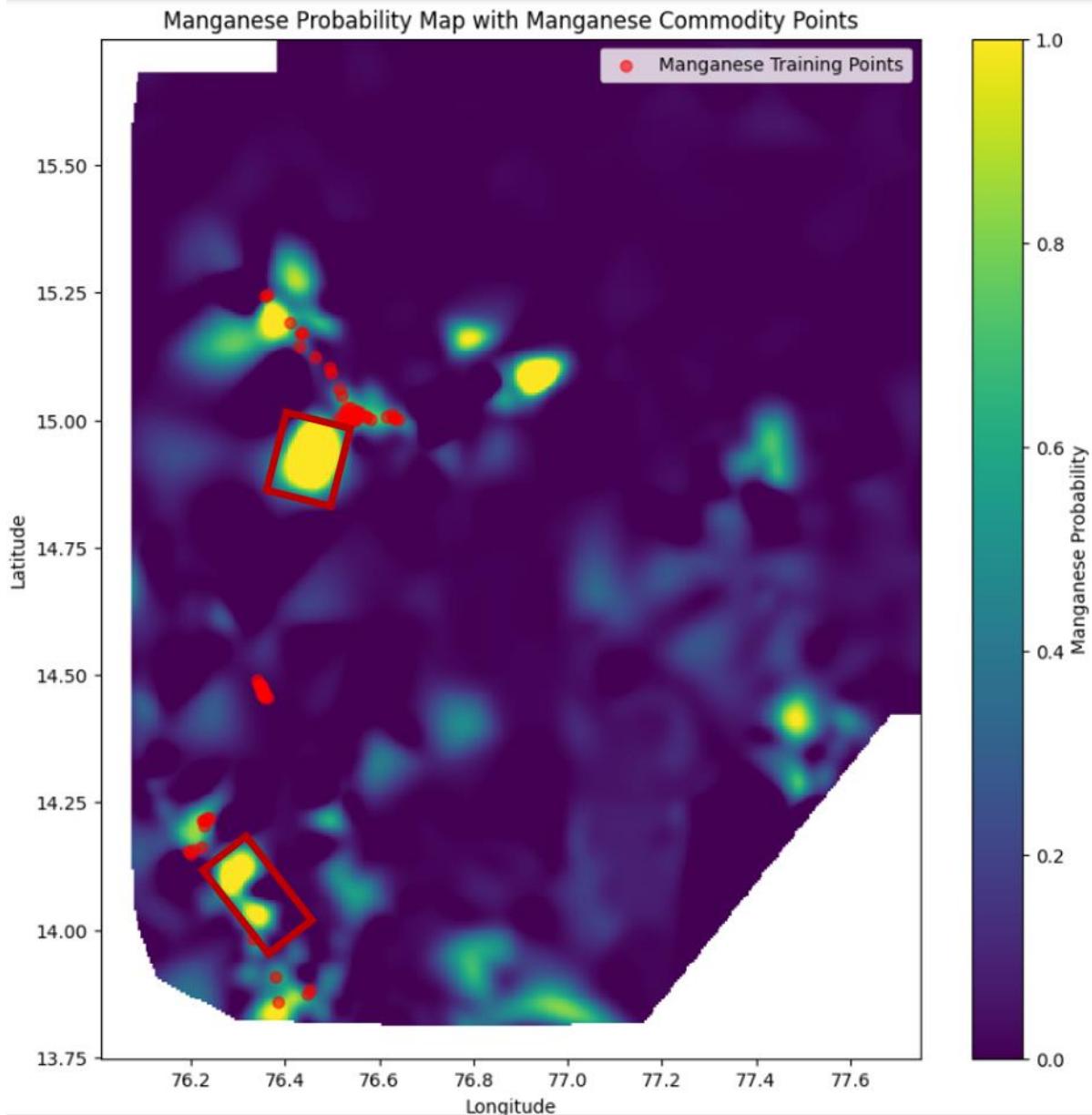


## Prospectivity Map

The prospectivity map generated from the ANN-1 architecture is based on the other prediction points which are not used in the model building. It is to be noted the ANN-1 structure input did not use the geological lithology and other information. So, the model is fully biased free from surface geology. The map shows there is a significant good prospective region exist where no commodity for Manganese is mentioned till date. These regions along with the known commodity region are taken for the further depth modelling.



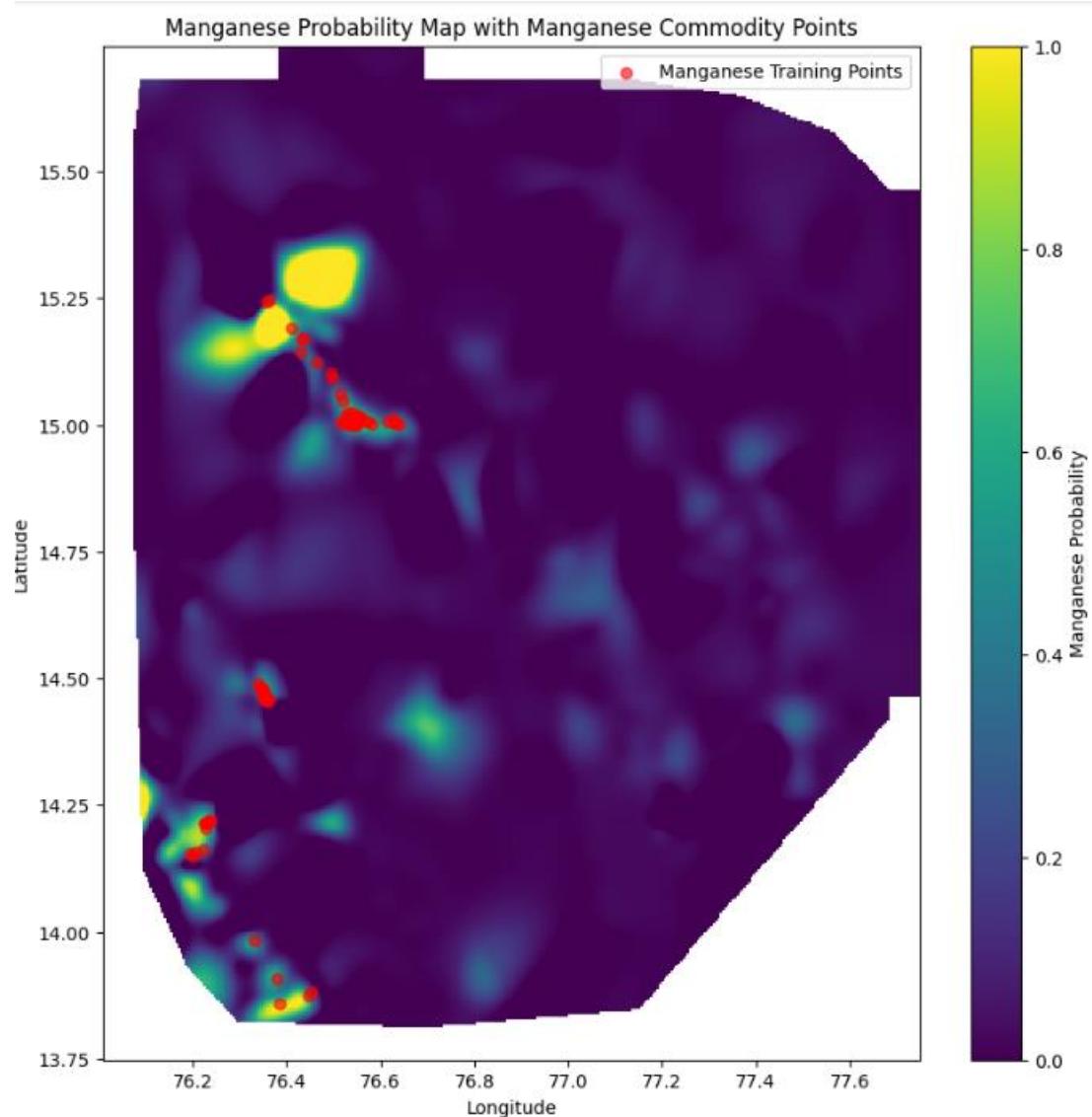
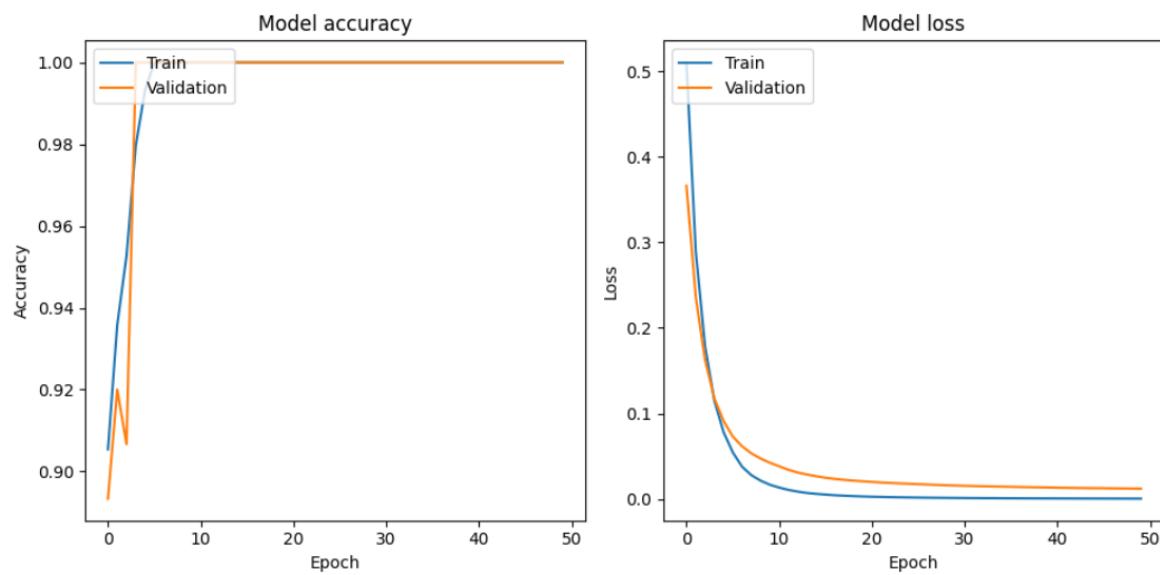
The marked area in the polygon with the red square can be taken for the further region of review as a result of the application of deep learning network:



## ANN-2

The ANN-2 deep-learning model is more complex with all the geological information along with the geophysical and geochemical information. As mentioned earlier the Categorical Data were converted to numerical data with one-hot encoding before the implementation of complex machine learning model

Training Accuracy	0.999
Validation Accuracy	0.9946
Training Error	0.00041
Validation Error	0.0012



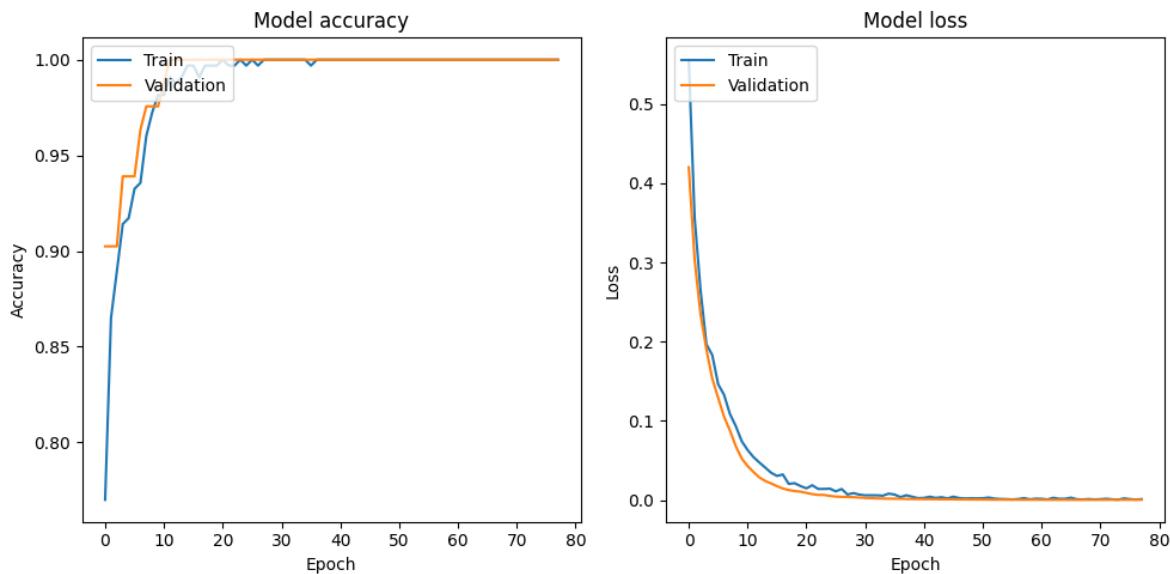
Prospectivity Map generated from ANN-2 Model

## Gold Model

### ANN-1

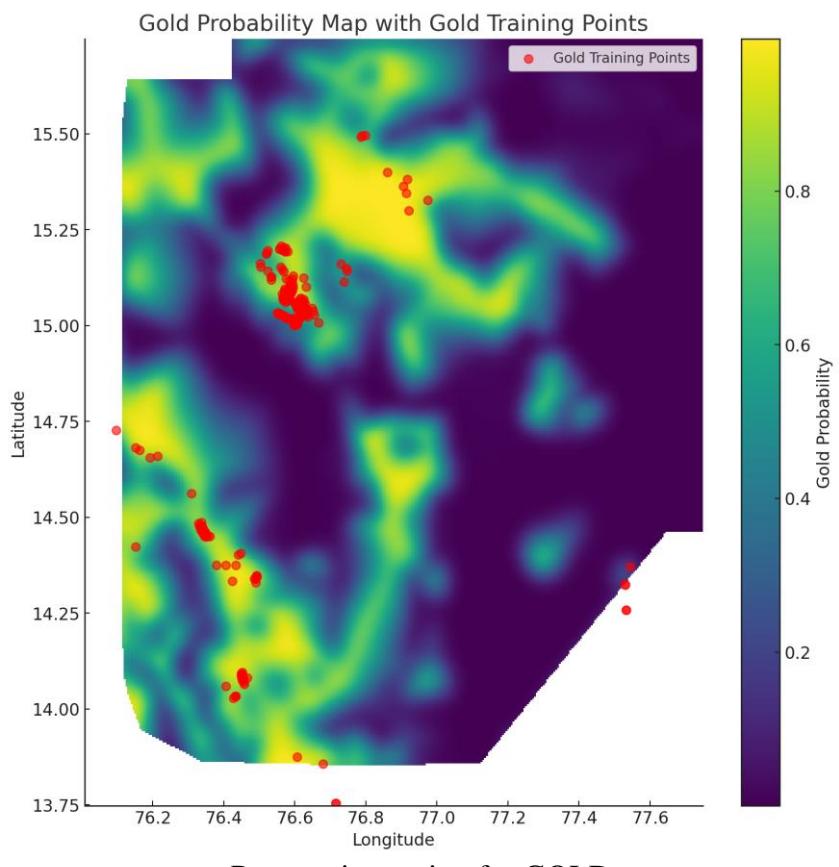
Training Accuracy	0.999
Validation Accuracy	0.999
Training Error	0.006
Validation Error	0.002

```
Epoch 2/100
11/11 0s 10ms/step - accuracy: 0.8596 - loss: 0.3873 - val_accuracy: 0.9024 - val_loss: 0.3032
Epoch 3/100
11/11 0s 10ms/step - accuracy: 0.8761 - loss: 0.2893 - val_accuracy: 0.9024 - val_loss: 0.2348
Epoch 4/100
11/11 0s 11ms/step - accuracy: 0.8980 - loss: 0.2168 - val_accuracy: 0.9390 - val_loss: 0.1901
Epoch 5/100
11/11 0s 11ms/step - accuracy: 0.8887 - loss: 0.2199 - val_accuracy: 0.9390 - val_loss: 0.1533
Epoch 6/100
11/11 0s 10ms/step - accuracy: 0.9328 - loss: 0.1560 - val_accuracy: 0.9390 - val_loss: 0.1291
```

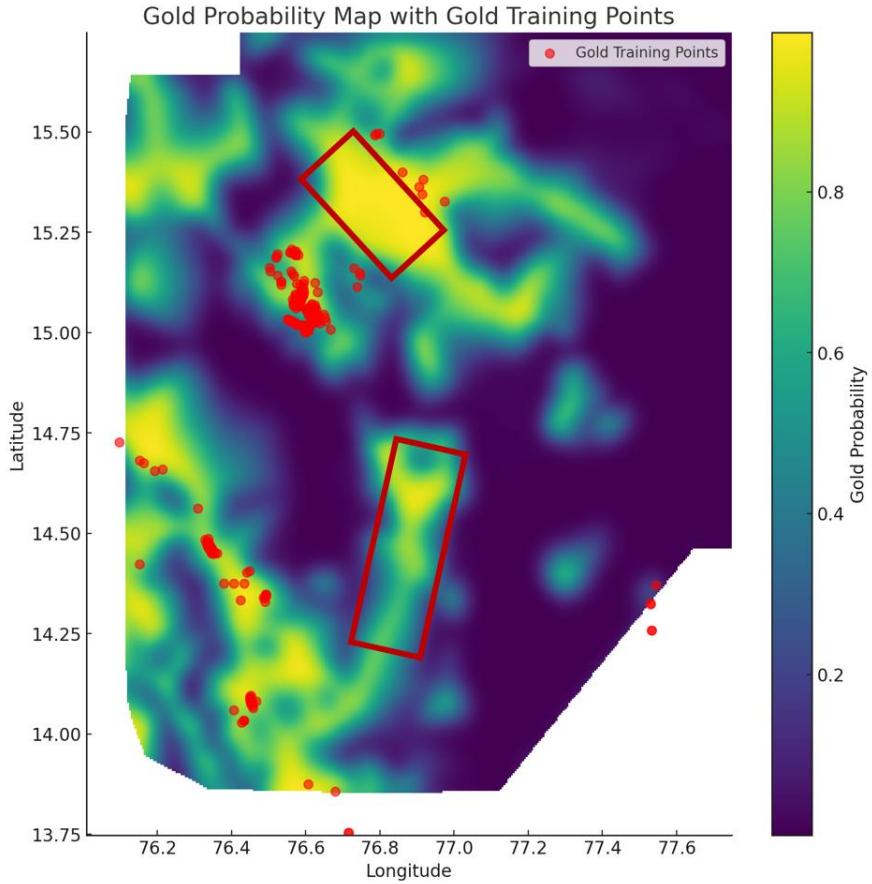


### Prospectivity Map

The Prospectivity map of ANN-1 as mentioned earlier does not take into the consideration of geological lithology and age input as provided in 50K data.



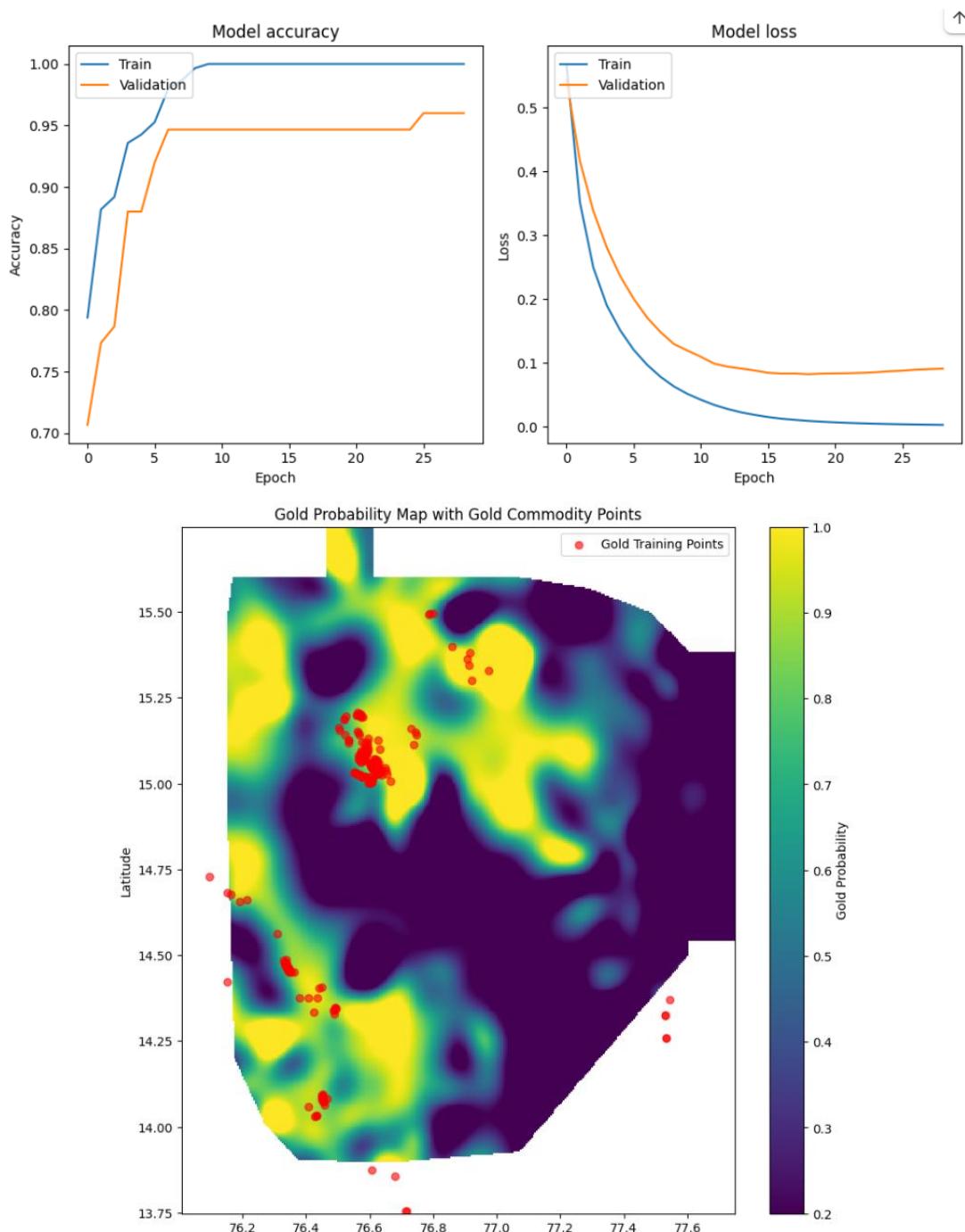
Prospective region for GOLD



## ANN-2

Inclusion of geological data creates more complexity and reduces the accuracy of the model. The training and validation error also little higher compare to the ANN-1 model.

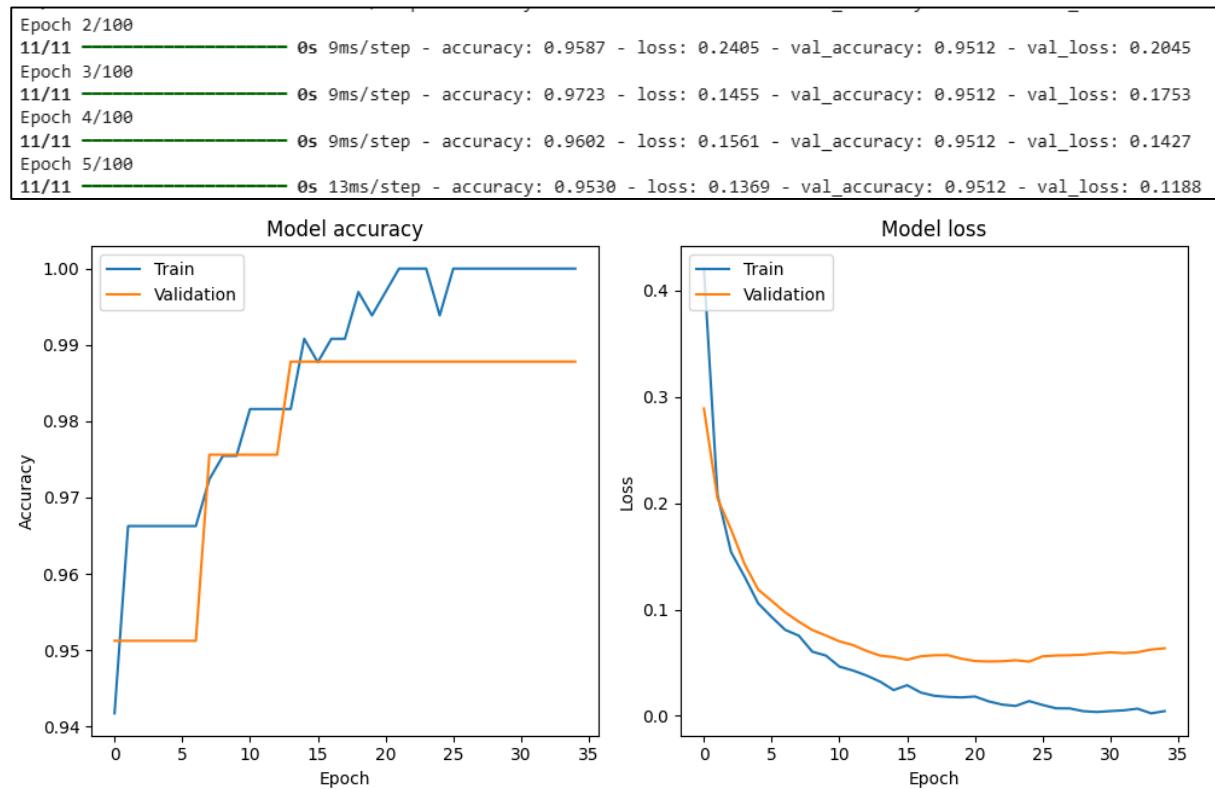
Training Accuracy	0.99
Validation Accuracy	0.94
Training Error	0.06
Validation Error	0.12



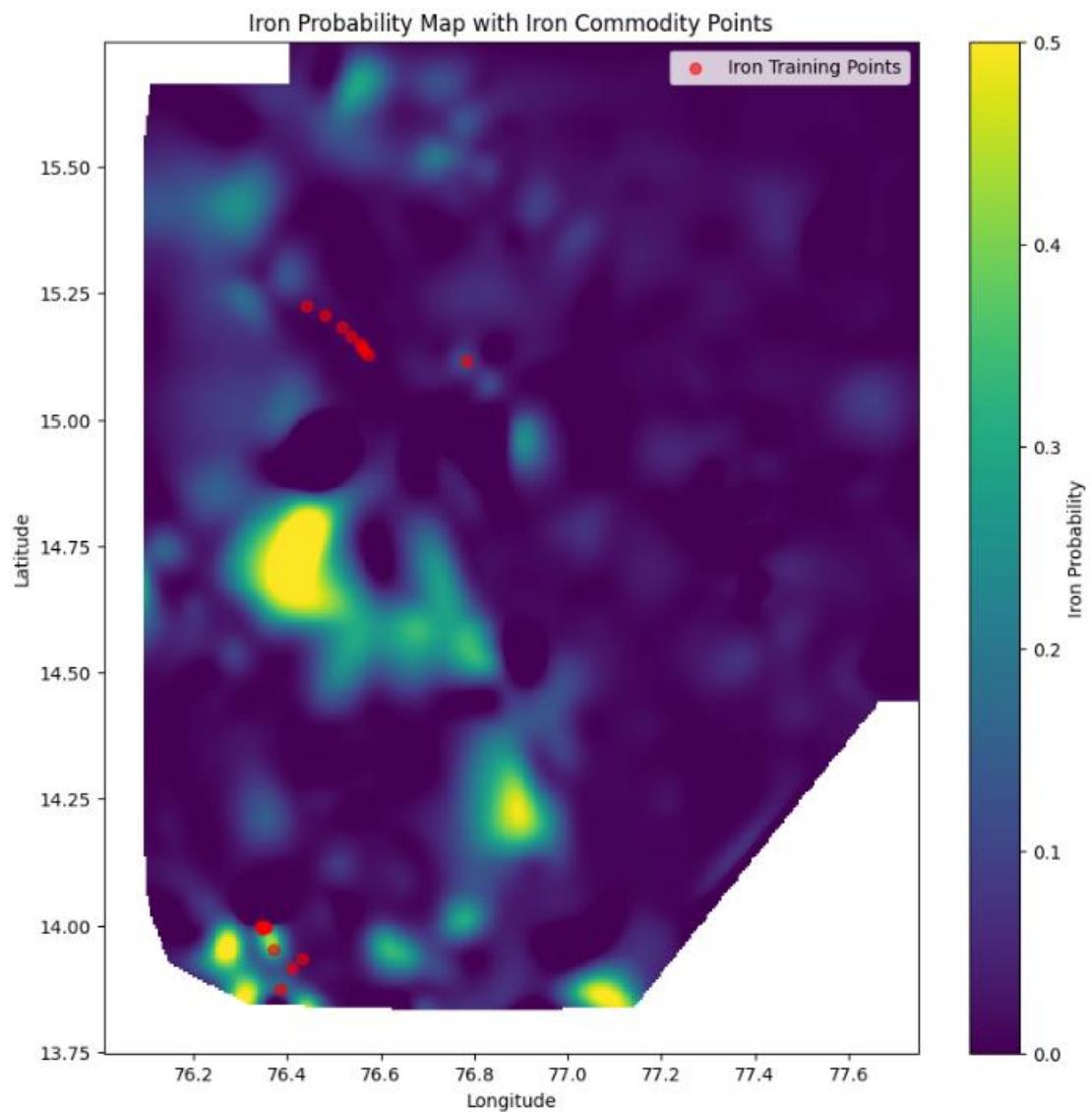
## Iron Model

### ANN-1

Training Accuracy	0.9926
Validation Accuracy	0.9878
Training Error	0.022
Validation Error	0.055



## Prospectivity Map



## UNSUPERVISED CLASSIFICATION OF THE FIELD DATA

Unsupervised clustering is a machine learning technique used to group data points into clusters based on similarity without prior knowledge of class labels. The goal is to identify patterns, structures, or relationships within the data. In this chapter we combine the power of unsupervised clustering with first *principles of geology for finding the target minerals* and commodities based on field geologic measurements.

### K-Means Clustering

The algorithm categorises ‘x’ observations into ‘z’ clusters ( $x \geq z$ ) for a set of observations S ( $S_1, S_2, \dots, S_n$ ) into a set B ( $B_1, B_2, \dots, B_z$ ) based on the Euclidean distance of the cluster centroids ( $\alpha$ ) as shown in the following equation:

$$\sum_{i=1}^x \sum_{x \in B_i} ||S - \alpha_i||^2$$

The process relies on iteratively refining the position of centroids to minimise the within-cluster variances, effectively grouping similar data points together. The recalculation of centroids based on the square of Euclidean distance between ‘ $B_i$ ’ and  $\alpha$  is estimated, and each observation is assigned the nearest calculated centroid as its respective cluster. The steps are repeated until centroids no longer change significantly indicating convergence.

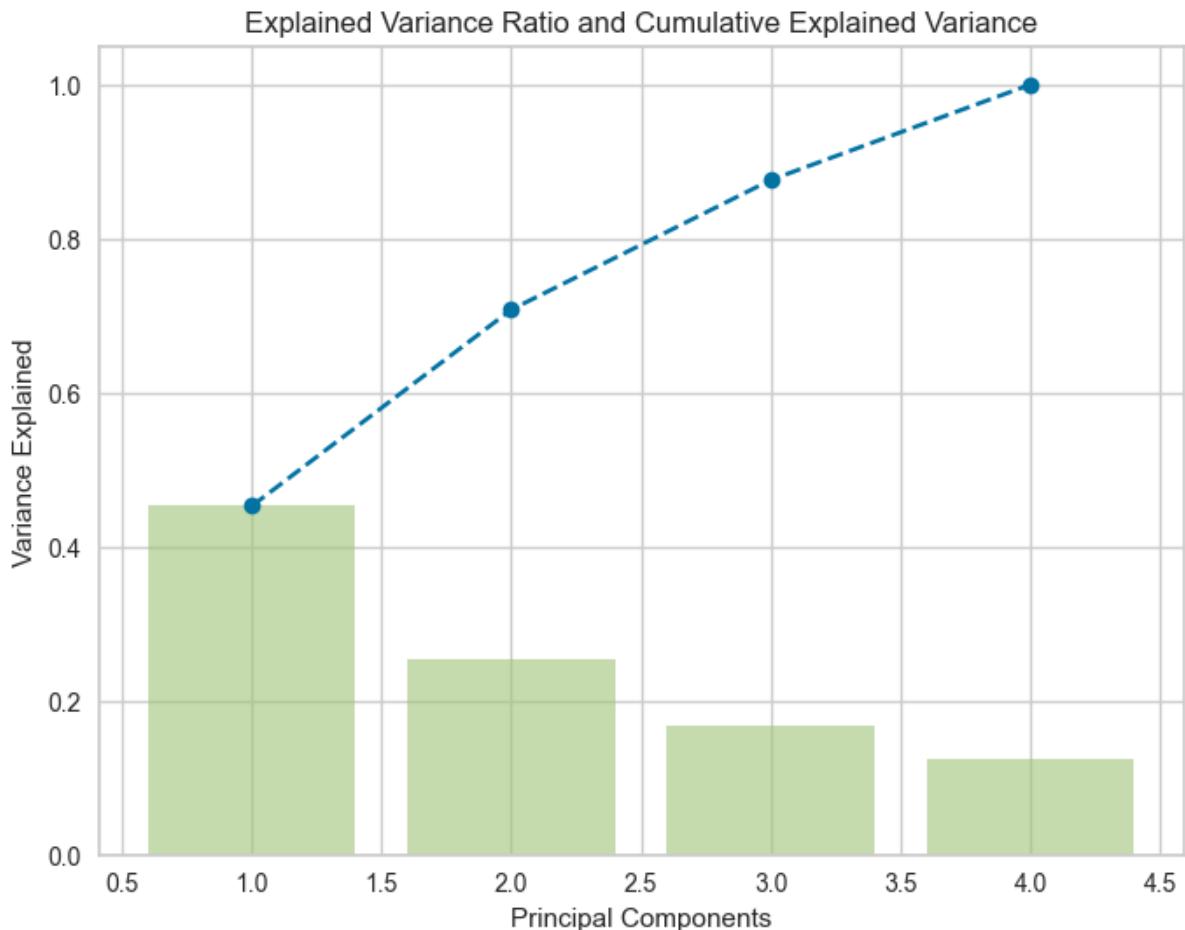
The optimum number of clusters is decided based on elbow and silhouette. In elbow method, the within cluster sum of squares (WCSS) is plotted against the number of clusters, where WCSS is a measure of variance within each cluster and calculated as sum of squares of centroid distances of clusters. The goal of clustering is to minimize WCSS, which would indicate that points within a cluster are as close to the centroid as possible, implying low variance within clusters.

Silhouette scores are a measure of how well a data point fits into its assigned cluster and is distinct from other clusters. It provides a quality of clustering and a measure of how internally coherent clusters and their separations are. The silhouette scores range between -1 and 1, with scores close to 1 indicating good clustering and -1 for poor clustering.

### Geological Data & Methodology:

The shape files present in the dataset ‘multi\_layer\_geological\_map\_of\_karnataka\_and\_andhra\_pradesh\_25k\_scale\_v1’ were first analysed and they consisted of data pertaining to dyke\_line, fault, fold, lithology, minerals, & oriented features. By combining all the shape files a common shape file was created and analysed for the clustering process. By analysing the data points and the attributes it was found that out of 16,039 data points there were four attributes namely ‘accessory\_’, ‘major\_mine’, ‘lithologic’, and ‘texture’ which comprised for 4,495 data points which was used in the

unsupervised classification process. Further a Principal Component Analysis of the 4 attributes was carried out to check for variance within datasets and all four contributed to significant variance with each other as illustrated below.



Principal Component Variance				
Principal Component	accessory_	major_mine	lithologic	texture
PC1	0.541704	0.614895	-0.146811	-0.553992
PC2	-0.167396	0.101559	0.934109	-0.298504
PC3	0.763076	-0.098829	0.323483	0.550733
PC4	-0.310242	0.775772	0.035288	0.548345

The above table and plot indicates the amount of variance each attribute brings to the 4 principal components. The principal components (PCs) are linear combinations of the original features. The loadings represent how much each original feature contributes to each principal component. The values in the table show the weight or importance of each feature in defining that principal component.

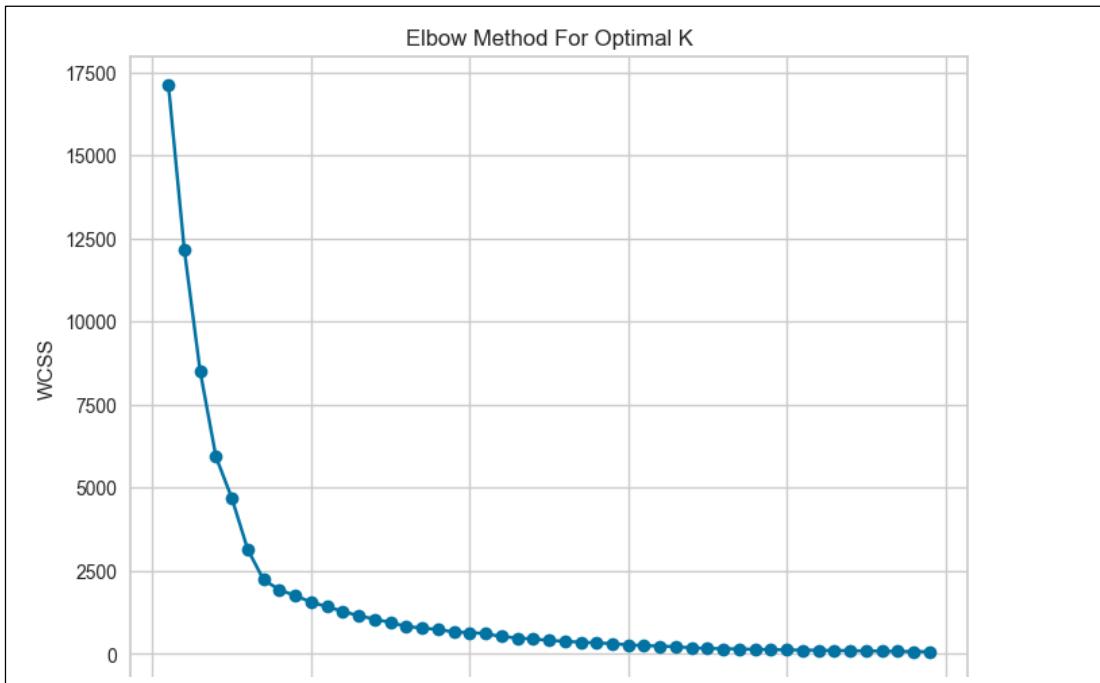
In short all the features play a very important role and have significant variance indicating mutually independent measurements.

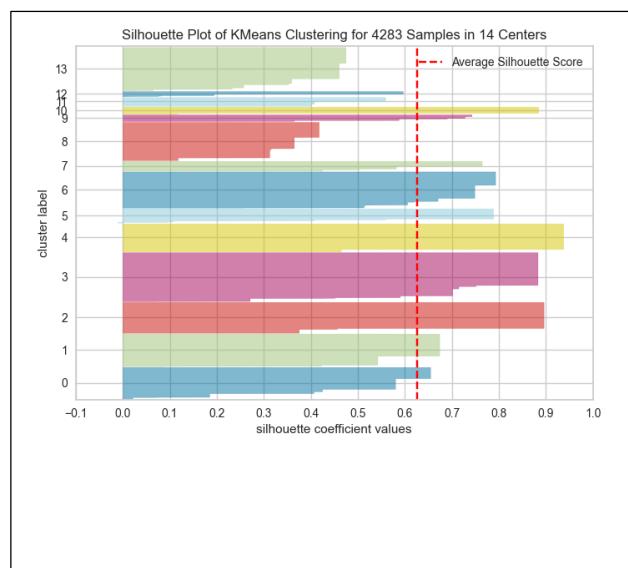
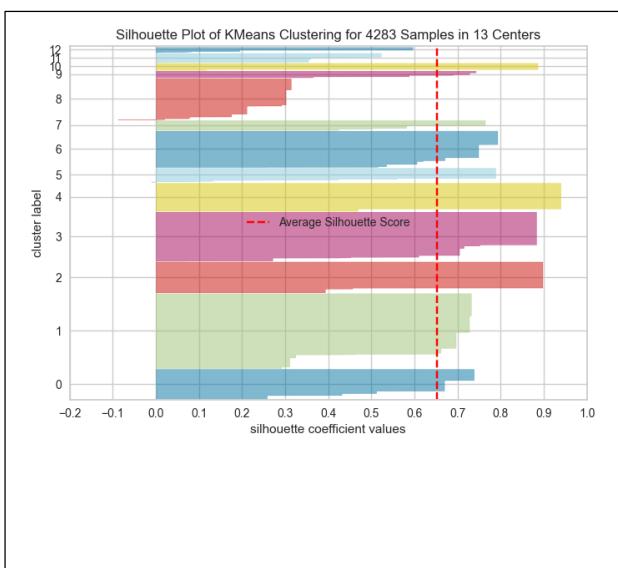
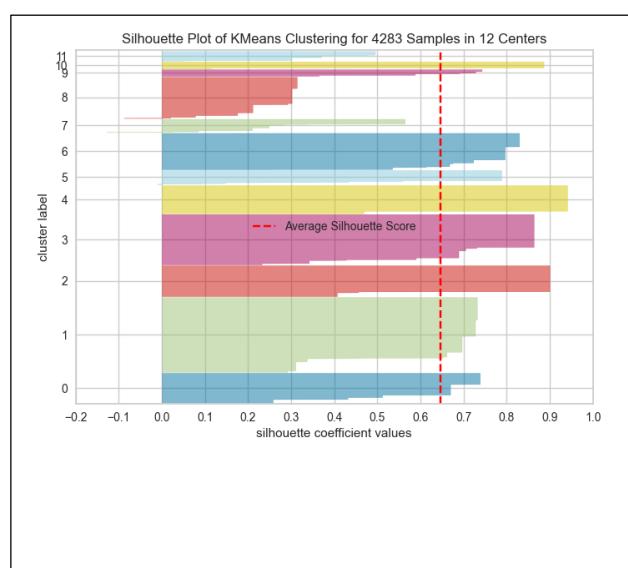
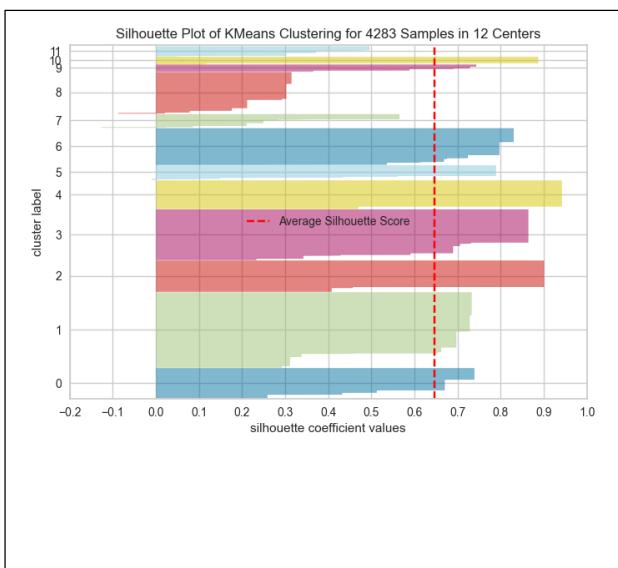
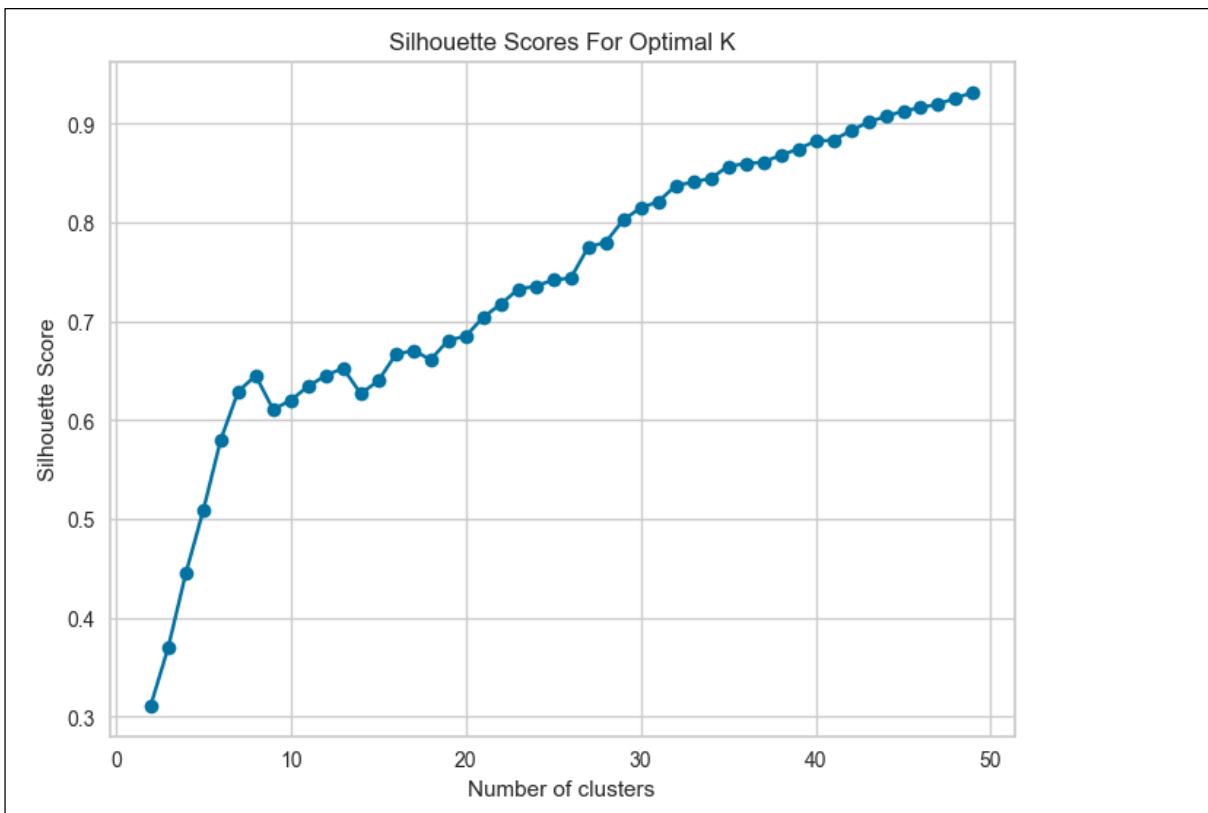
The geological information present in the shape files was simplified based on the following table for clustering purposes with cluster number denoted beside them.

Classification of Geological Field Attributes			
Accessory Minerals (accessory_)	Major Minerals (major_mine)	Lithology (lithologic)	Texture (texture)
1-> Opaques and Iron Oxides	1-> Mica Group	1-> Gabbro / Dolerite-related	1-> Foliated or Layered
2-> Silicate Minerals and Associated Group	2-> Quartz Group	2-> Granite / Gneiss-related	2-> Fine to Coarse Grain
3-> Quartz-related and Silica-rich	3-> Mica Group	3-> Quartz / Quartzite-related	3-> Porphyritic and Xenoblastic Textures
4-> Calcite and Carbonates	4-> Amphibole Group	4-> Schist and Phyllite-related	4-> Vesicular and Pitted Textures
5-> Phosphate and Rare Earth Elements	5-> Pyroxene Group	5-> Pegmatite-related	5-> Ophitic and Spinifex Textures
6-> Titanium and Zirconium-related	6-> Chlorite Group	6-> Conglomerates	6-> Clastic Textures
7-> Others ( dunite-peridotite-pyroxenite and fluorite and sphene)	7-> Carbonates	7-> Others (Meta ultramafite and MS Suite)	7-> Mosaic and Ribbon Textures
	8-> Iron and Manganese Minerals	8-> Metavolcanic/volcaniclastic	8-> Hypidiomorphic, Granular, and Equigranular Textures
		9-> Iron-rich & pyroclastic	
		10-> Mafic-ultramafic intrusives	
		11-> Carbonate rocks	

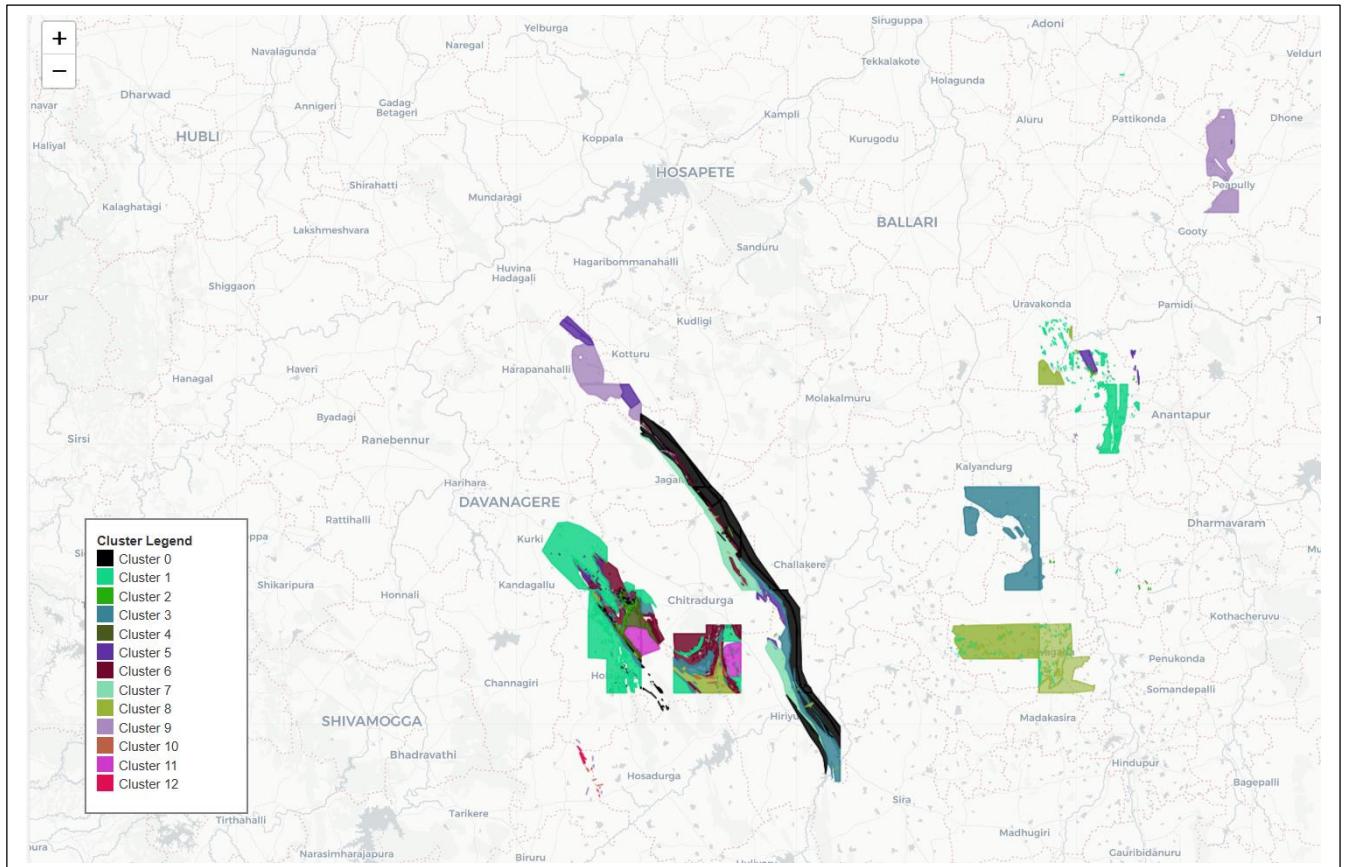
All the dataset used in the clustering process is depicted and shown in the following figure. The areas shaded in purple colour depicts the study area where all the data points for the four attributes used in the classification process are present without any <null> values. The maps are interactive in nature and are plotted using the Folium Library in Python and can be saved in .html format.

Before the clustering process the Elbow and Silhouette Plots were generated for 1-50 clusters in order to find out the optimum number of clusters as depicted below. Based on the analysis of the plots it can be seen that at 13 number of clusters there is a very gentle rise in slope in the Silhouette Plot and also the Silhouette scores are higher than 14 clusters. The Elbow plot also corroborates with change in slope at 13 number of clusters.





Based on the results of clustering the following map is generated once again in Folium which can be accessed while running the code. The table after that summarises the findings of the results.



Cluster	Accessory Minerals (accessory_)	Major Minerals (major_mine)	Lithology (lithologic)	Texture (texture)	Target Commodities & Minerals
0	Silicate Minerals and Associated Group	Quartz Group	Granite / Gneiss-related	Foliated or Layered	<b>Gold, Copper</b>
1	Quartz-related and Silica-rich	Feldspar Group	Granite / Gneiss-related	Hypidiomorphic, Granular, and Equigranular Textures	<b>Copper</b>
2	dunite-peridotite-pyroxenite or fluorite and sphene	Iron and Manganese Minerals	Quartz / Quartzite-related	Foliated or Layered	<b>Iron, Mn</b>
3	Quartz-related and Silica-rich	Feldspar Group	Meta ultramafite or MS Suite	Weathering & Fracture Textures	<b>Ni, Copper</b>

4	Silicate Minerals and Associated Group	Iron and Manganese Minerals	Schist and Phyllite-related	Foliated or Layered	<b>Iron, Mn</b>
5	dunite-peridotite-pyroxenite or fluorite and sphene	Mica Group	Granite / Gneiss-related	Foliated or Layered	<b>Ni-PGE, Copper</b>
6	Silicate Minerals and Associated Group	Feldspar Group	Meta ultramafite or MS Suite	Fine to Coarse Grain	<b>Ni-PGE, Copper</b>
7	Opaques and Iron Oxides	Chlorite Group	Meta ultramafite or MS Suite	Porphyritic and Xenoblastic Textures	<b>Ni</b>
8	Silicate Minerals and Associated Group	Mica Group	Granite / Gneiss-related	Porphyritic and Xenoblastic Textures	<b>Iron, Mn</b>
9	dunite-peridotite-pyroxenite or fluorite and sphene	Mica Group	Granite / Gneiss-related	Hypidiomorphic, Granular, and Equigranular Textures	<b>Ni, Copper</b>
10	Calcite and Carbonates	Feldspar Group	Gabbro / Dolerite-related	Ophitic and Spinifex Textures	<b>Ni-PGE, Copper</b>
11	Calcite and Carbonates	Carbonates	Schist and Phyllite-related	Vesicular and Pitted Textures	<b>REE</b>
12	Phosphate and Rare Earth Elements	Amphibole Group	Meta ultramafite or MS Suite	Weathering & Fracture Textures	<b>REE</b>

### **Discussion:**

The power of K-Means clustering is demonstrated by transforming field geological data into clusters which represent areas where targeted minerals and commodities such as REE, Ni-PGE, Copper, Iron, Manganese, and Gold can be found. The interpretation of the clusters is based on the interpretation of the combination of geological attributes. In this method the first principles of geology are combined with results derived from clustering in order to derive meaningful information pertaining to finding minerals & commodities within the cluster. It may be noticed that for some clusters such as 10, 6, & 5 Ni-PGE & Copper are likely to be found. All the 3 clusters owe their metal enrichment to an ultramafic-derived sulfide source, but they trap and preserve those sulfides in very different geological settings and therefore differ in host rock, trapping mechanism and resulting ore geometry.

In cluster 10 it is found in a layered mafic intrusion or possibly komatiite flow. Often high-grade, stratiform or discordant pods are found in the feeder zones. In cluster 5 Felsic (granite/gneiss) carrying ultramafic slivers where primary magmatic sulfide layer preserved through regional metamorphism and locally sheared into pods along foliation. They can offer moderate-tonnage orebodies. In cluster 6 the host rock is a Metamorphosed ultramafic flows/intrusions, and can offer high grade zones in small aerial pockets.

Similarly, in clusters 2, 4, & 8 offer Iron & Mn as the target commodity but the difference being their occurrence and genesis. In cluster 2 the genesis is related to Hydrothermal fluids from an intruding ultramafic body which reacts with silica-rich quartzite, precipitating massive Fe–Mn silicates and oxides in a stratabound skarn layer. They typically yield large tabular feeder zones and are extensively present. In cluster 4 their genesis is related to chemical sedimentary (exhalative) Fe–Mn layers that were then deformed and recrystallized into schist/phyllite. They also extend large distances but their thicknesses are smaller and discontinuous as compared with cluster 2. In cluster 8 the genesis is related to late-stage, metal-bearing fluids circulating through fractures in granite while depositing oxides and carbonates in veins. They typically extend short distances but can have high grades within their pockets of occurrence. Similarly, each cluster has its unique geological history and genesis which present a unique opportunity to prospect for corresponding target commodities and minerals as indicated.

## DEPTH MODELING FOR MINERALISED BODIES

### 1. Introduction

The cause from the set of observations is primarily provided by geophysical inversion. The distribution of physical characteristics (such as density, conductivity, magnetic susceptibility, etc.) beneath the surface of the Earth is what causes the geophysical anomalies that are seen there. Geophysical inversion is the final stage of geophysical interpretation, which aids the interpreter in comprehending the geometry, depth, and type of rock found in the various formations within the research area. The presence of the forward solution is a prerequisite for the existence of an inverse solution.

There is no depth resolution available in the potential field data. Therefore, the data must be inverted in order to retrieve the right depth information. A mathematical framework for obtaining a trustworthy model of the earth's subsurface from the measured data sets is provided by the inversion of these data (density for gravity data and susceptibility for magnetic data, for example). Inverse problems can be broadly classified into two categories: generic inverse problems (underdetermined) occur when the number of unknowns exceeds the number of data points, and parameter estimation (overdetermined) problems occur when the number of data points exceeds the number of unknowns.

This chapter represents a 3D gravity forward modeling and inversion algorithm for finding out the density distribution at different depths below the surface. The model used here is a discretized prism model and inversion technique incorporates gradient descent algorithm.

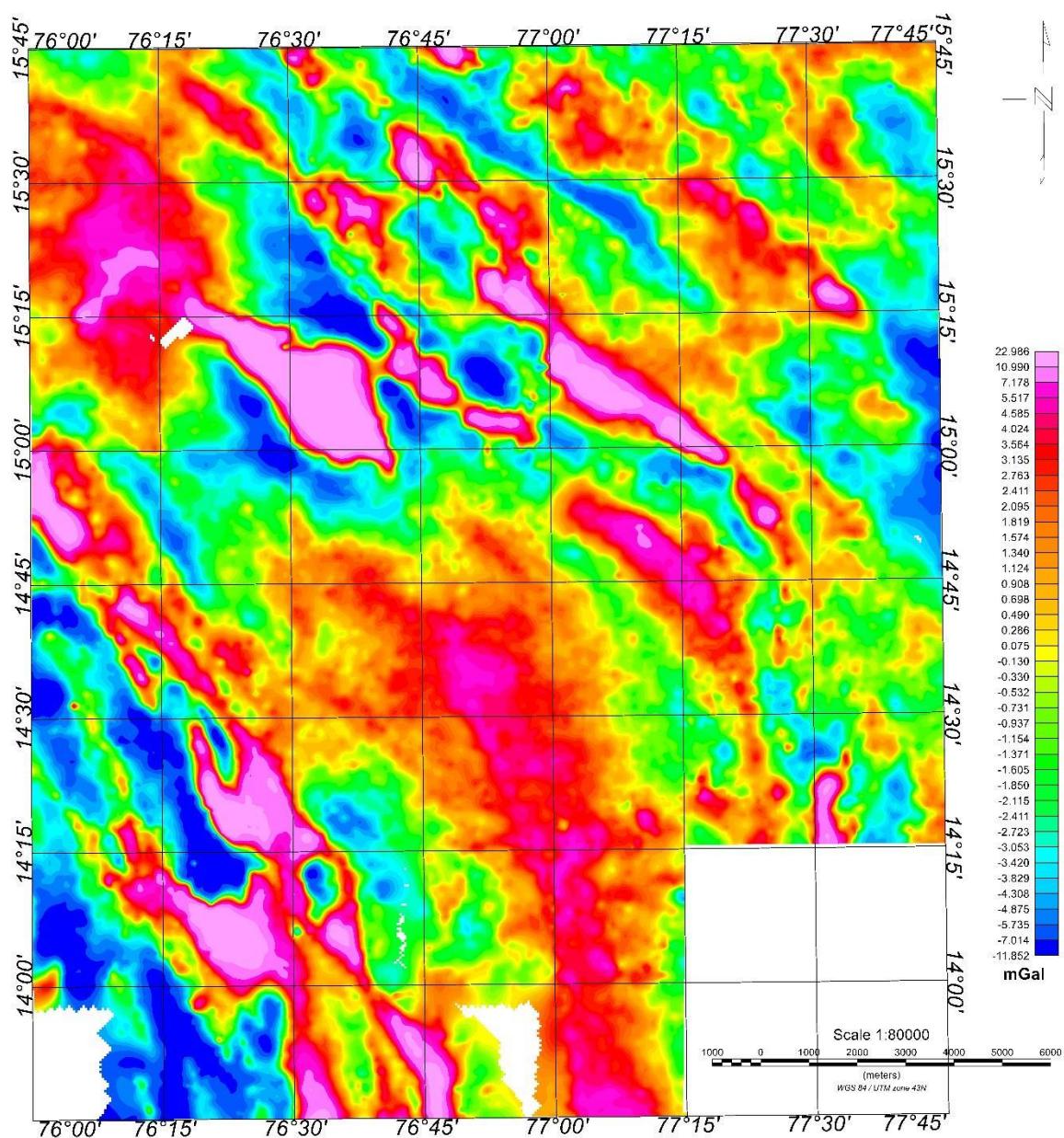
## 2. Data Preparation

### 2.1 Data used for the model preparation

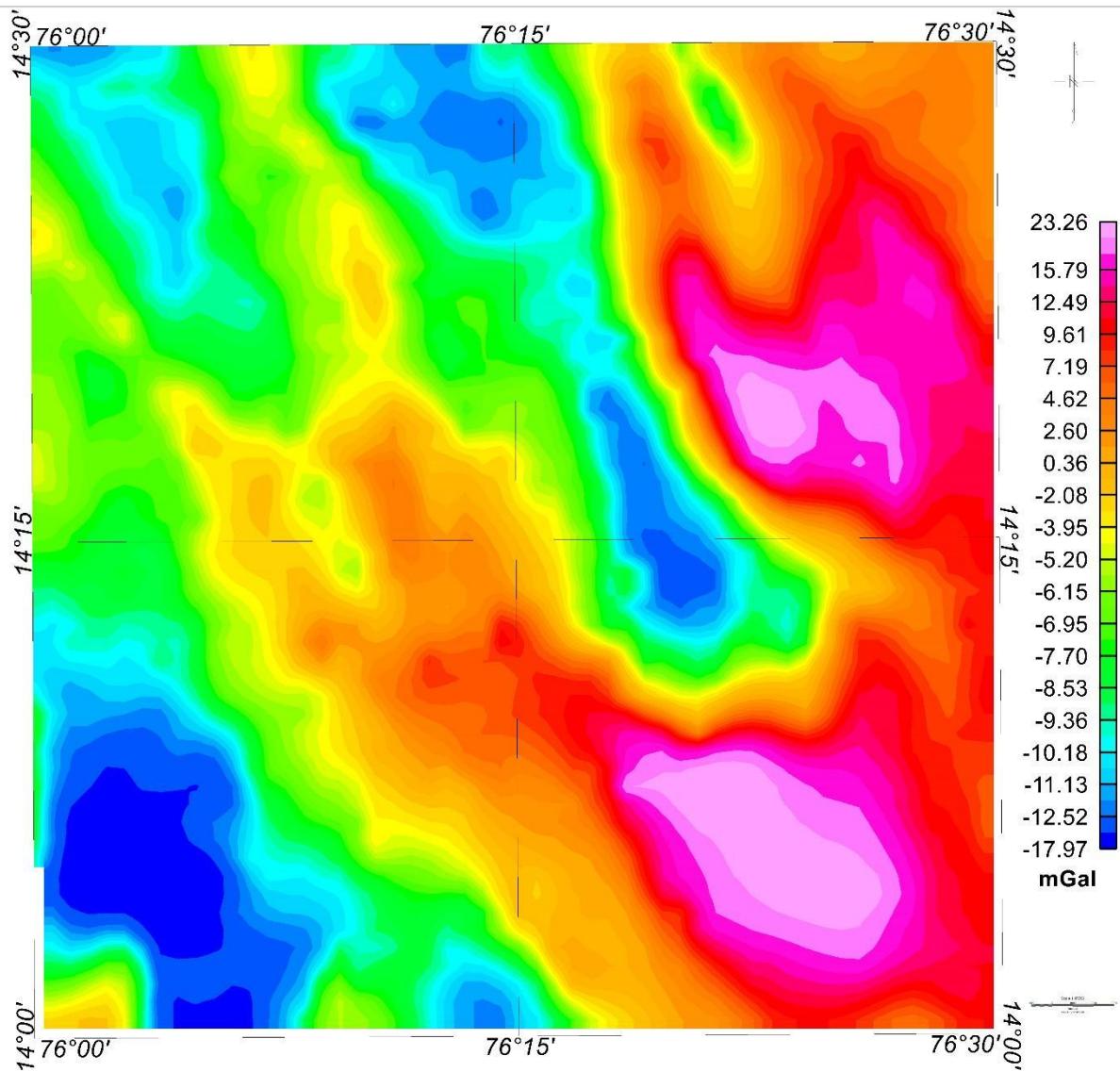
The Bouguer gravity anomaly map represents the combined effect of the regional and residual anomalies lying at different depths in the subsurface. So, we have used the provided Bouguer anomaly grid and obtain the Residual gravity grid. Again, we have re-grid the Residual Bouguer anomaly grid to a grid spacing of 1000m. The aim to re-grid the data is to interpolate the irregularly spaced provided gravity data observations onto a regular grid to perform the 3D inversion.

As the area is very big so, we have only considered the zone of interest (**parts of Chitradurga Schist belt where high residual BA anomaly exists and gold and Iron commodity present**) for calculating the depth. The study area is bounded by latitudes 14°00'N to 14°30'N and 76°00'E to 76°30'E. The total number of interpolated points used for the inversion is 2961.

## RESIDUAL BOUGUER ANOMALY MAP OF KARNATAK AND ANDHRA PRADESH



## PARTS OF RESIDUAL BOUGUER GRAVITY DATA EXTRACTED FROM RESIDUAL ANOMALY MAP AND IS USED FOR DEPTH MODELING



### 2.2 Prism Grid Definition

The area is divided into rectangular prisms for applying the 3D inversion. The extension of the study area in X-direction is 54km and in Y-direction is 55km. The grid cell size used for preparing the model in X-direction is 1000m and Y-direction is 1000m. The depth range is from 0 to 3km with a grid spacing of 200m. So, the total nos. of prisms used are ( $54 \times 55 \times 15 = 44550$ ) and the no. of observed data points are 2961.

### 3. Forward modeling Formulation

#### 3.1 Initialization of Density Model

We have also used the Physical property datasheet and used the density measurement values of the rock samples for initializing the model. The main lithology used for the model is as follows:

(i) **Hornblende Granite having density value 2.75gm/c** (ii) **Pink granite having density value 2.58 gm/cc**, and (iii) **BIF with quartz epidote vein having density value 3.1gm/cc.**

### 3.2 Prism Boundaries and prism centers

Boundaries in X, Y, and Z directions were computed and centers of each prism were calculated using python.

### 3.3 Sensitivity matrix calculation

The mass of the Earth generates its gravitational field. Therefore, by integrating "g" everywhere at the surface, the flux of the gravity field at any given surface can be found. This flux is proportional to the mass "M" that the surface encloses.

Mathematically this can be expressed

$$\oint g \cdot ds = 4\pi M = 4\pi G \int \sigma dV \quad (1)$$

Where G is the universal gravitational constant.

The gravity field g is curl free hence is a conservative field.

$$\text{So, } g(r) = \nabla U \quad (2)$$

Using Gauss's theorem, the surface integral can be converted to volume integral. So, Equation (1) can be written as

$$\oint_S g \cdot ds = \int_V \nabla \cdot g dV = \int \nabla^2 U dV \quad (3)$$

Comparing Equation (1) and (3), it can be written as

$$\nabla^2 U = 4\pi G \sigma \quad (4)$$

The potential satisfy Dirichlet boundary condition at infinity

$$\text{i.e., } U = 0 \text{ at } r = \infty$$

For problems of geophysics, the observation points stay outside the mass hence net gravitational field flux becomes zero.

The gravitational potential will satisfy Laplace's Equation at these observation points.

$$\text{i.e., } \nabla^2 U = 0 \quad (5)$$

### Integral representation of gravity anomaly

The vertical component of the gravitational attraction of rectangular prism with dimensions  $x_1$ , to  $x_2$ ,  $y_1$  to  $y_2$  and  $z_1$ , to  $z_2$ , at the observation point  $(0, 0)$  is expressed as

$$\delta g_k = -G\sigma_k \int_{x_1}^{x_2} dx' \int_{y_1}^{y_2} dy' \int_{z_1}^{z_2} dz' \frac{z'}{(x'^2 + y'^2 + z'^2)^{3/2}}. \quad (6)$$

where  $G$  is the universal gravitational constant.  $x'$ ,  $y'$  and  $z'$  is the distance between the center of the rectangular prism and observation point  $(0, 0)$  in the x, y, and z directions, respectively. Blackly (1996) provided the solution of Equation (6) as

$$\delta g_k = -G\sigma_k \sum_{\alpha=1}^2 \sum_{\beta=1}^2 \sum_{\gamma=1}^2 \varsigma_{\alpha\beta\gamma} \left[ z_\gamma \tan^{-1} \left( \frac{x_\alpha y_\beta}{z_\gamma R_{\alpha\beta\gamma}} \right) - x_\alpha \ln(R_{\alpha\beta\gamma} + y_\beta) - y_\beta \ln(R_{ijk} + x_\alpha) \right]. \quad (7)$$

where  $\delta g_k$  is the vertical component of the gravitational attraction due to the  $k^{\text{th}}$  rectangular prism,  $\sigma_k$  is the density of the  $k^{\text{th}}$  rectangular prism,  $R_{\alpha\beta\gamma} = \sqrt{x_\alpha^2 + y_\beta^2 + z_\gamma^2}$ , and  $\varsigma_{\alpha\beta\gamma} = (-1)^\alpha (-1)^\beta (-1)^\gamma$ . Total gravitational field due to all rectangular prisms can be written as the summation of field due to individual prism at observation points  $(0, 0)$ , Thus:

$$g_z = \sum_{k=1}^M \delta g_k \quad (8)$$

where,  $M$  is the total number of rectangular prisms.

Let  $(x_0^i, y_0^i)$  ( $i = 1, 2, 3, \dots, N$ ) be the coordinates of the data points where  $N$  is the number of data points in the survey area. Then, the vertical component of the gravity field at the  $i^{\text{th}}$  data point is given by

$$d_i = \sum_{k=1}^M a_{ik} \sigma_k \quad (9)$$

where,  $\sigma_k$  is the density of the  $k^{\text{th}}$  rectangular prism and  $a_{ik}$  matrix element representing the influence of the block on the data point and can be formulated using Equation 9. The Equation (9) can be written in matrix notation

$$\mathbf{d}^0 = \mathbf{A} \mathbf{m} \quad (10)$$

where  $\mathbf{d}^0$  is the observed vertical component of the gravity,  $\mathbf{m}$  is the model parameter vector and  **$A$  is the dimensional sensitivity matrix** and  $N$  and  $M$  are the number of observed data and total number of rectangular prisms, respectively.

### Vectorization and Parallel computing

Matrix formulation (equation 10), as demonstrated in the preceding section, entails three processes. 1) Calculation of gravity response (i.e.,  $g$ ) at a certain position caused by a single rectangular prism. 2) Using equation (9), the overall gravity field is calculated by adding together all of the individual rectangular prisms algebraically. 3) Steps 1 and 2 are repeated at other observation points. It is quicker for small-scale issues. However, the number of rectangular prisms and observation points in large-scale 3D gravity modeling range in the thousands to millions. As a result, the forward formulation will require a significant amount of calculation time. Additionally, as detailed in Wang et al. (2015), parallel computing is used to speed up forward commutation.

### 3.4 Gravity Prediction

The predicted gravity at each observation point was calculated using above approach and initial residual and RMS misfit between observed and calculated gravity were computed.

## 4. Inversion Methodology

Conjugate Gradient Descent (CGD) is a potent iterative optimization technique for effectively resolving large-scale linear inverse problems. CGD is especially well-suited for obtaining subsurface density distributions from gravity anomaly data in the context of geophysical inversion. This study describes how the CGD approach is implemented, its mathematical foundation, and its efficacy in our 3D gravity inversion algorithm.

### 4.1 Objective

The primary objective of the inversion is to use a forward model as represented by the sensitivity matrix to estimate the subsurface density contrast model that best fits the observed gravity anomaly data.

The forward problem is represented as:

$$d = G * m$$

Where:

d: observed gravity data (e.g., Bouguer anomaly)

G: sensitivity matrix (relating model parameters to data)

m: model vector (density contrast of each prism)

The inverse problem attempts to recover m given d and G.

### 4.2 Conjugate Gradient Algorithm Overview

The CG approach iteratively refines a model by traveling along conjugate directions, greatly decreasing compute and memory demands, in contrast to direct matrix inversion, which is computationally costly for large systems.

Steps of the Algorithm:

#### 1. Initialization:

- Start with an initial model  $m_0$  (often a zero or prior model).
- Compute the initial residual:  $r_0 = d - G * m_0$
- Compute initial gradient:  $g_0 = -G^T * r_0$
- Set search direction:  $p_0 = -g_0$

#### 2. Iteration:

For each iteration k:

- Compute step size  $\alpha_k$ :  $\alpha_k = (g_k^T * g_k) / (p_k^T * G^T * G * p_k)$
- Update model:  $m_{k+1} = m_k + \alpha_k * p_k$
- Update residual:  $r_{k+1} = r_k + \alpha_k * G * p_k$
- Compute new gradient:  $g_{k+1} = -G^T * r_{k+1}$
- Compute Fletcher–Reeves scalar:  $\beta_k = (g_{k+1}^T * g_{k+1}) / (g_k^T * g_k)$
- Update search direction:  $p_{k+1} = -g_{k+1} + \beta_k * p_k$

### 3. Stopping Criteria:

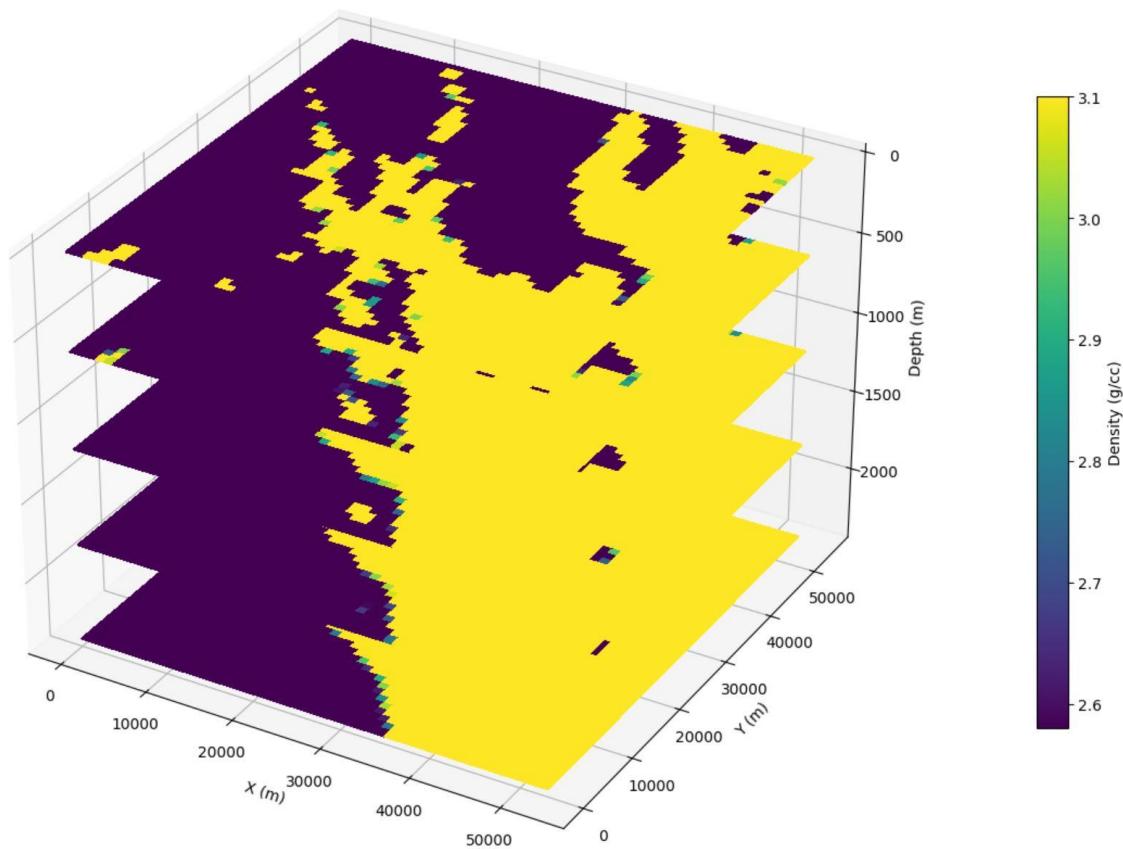
- Stop when residual norm or RMS error falls below tolerance (e.g., 0.5 mGal).
- Alternatively, limit iterations to prevent overfitting.

## Results and Discussions

The Root Mean Square (rms) error was calculated for the above model and the misfit error is nearly equal to 0.5. Further modification of the model can be achieved by using other inversion methods like data-space inversion technique. Using the forward modelling and inversion technique 3D horizontal depth slices plot was prepared.

- 3D horizontal plot shows that the parts of Chitradurga Schist belt where high residual BA anomaly exists and gold and Iron commodity present, is of greater depth more than 3km.

3D Horizontal Slices of Inverted Density Model



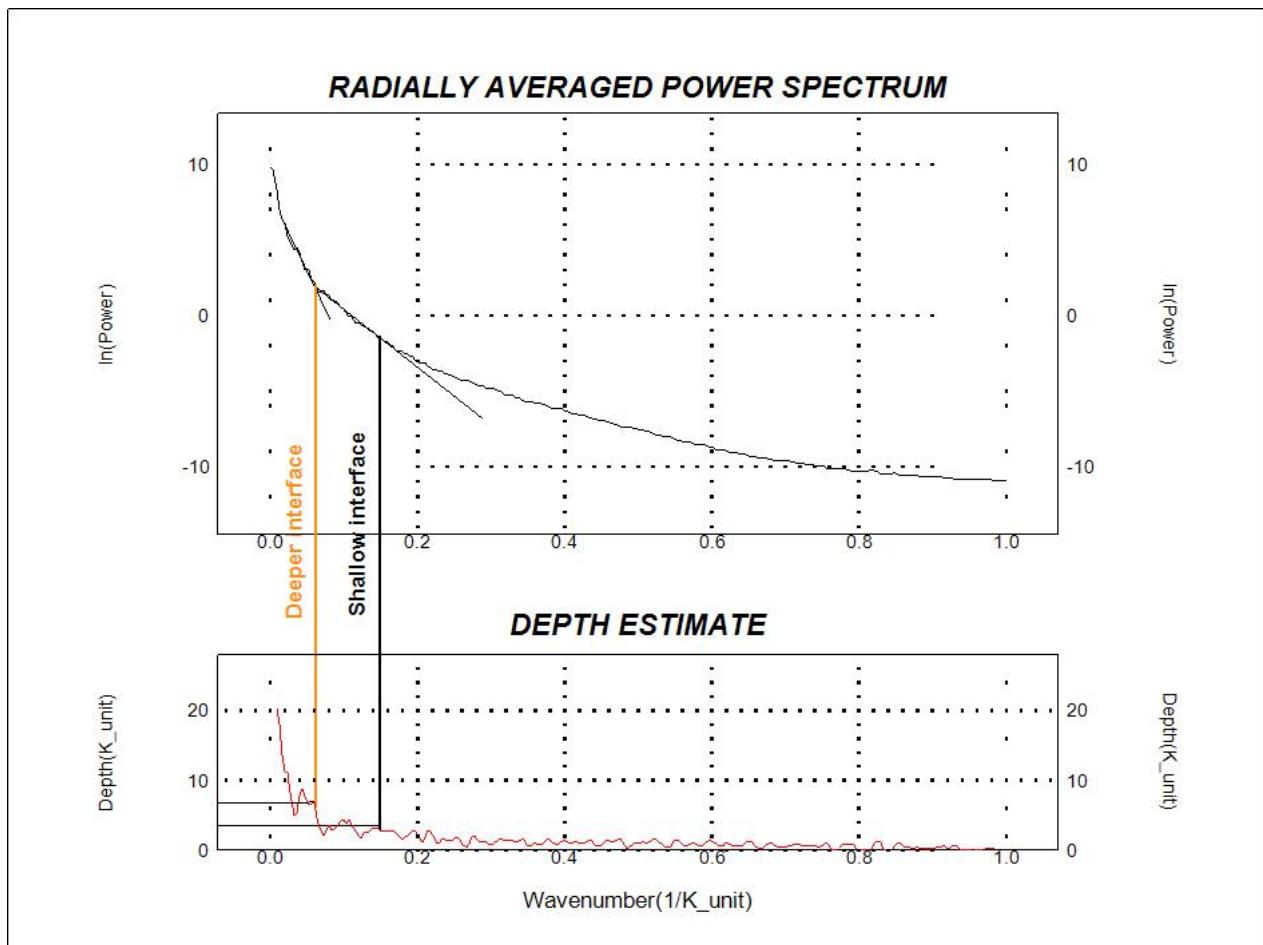
## Radially averaged power spectrum of Gravity anomalies

Spectral analysis of gravity and magnetic data is a conventional technique for the determination of depths of geological features, particularly the basement (Spector and Grant, 1970; Maus and Dimri, 1996). Spectrum segments with smaller wave numbers correspond to deeper features whereas higher wave numbers represent shallow features. Slopes of spectral

segments yield the estimates of average depths to the source bodies. The depth of each ensemble of sources responsible for each segment of a power spectrum is computed from the following relation:

$$\text{Depth} = -\text{Slope}/4\pi$$

In the present case, the radially averaged power spectra of gravity anomaly has been computed for the study area, which shows two spectrum segments for which regression lines are fitted. Slopes of spectral segments correspond to average ensemble depths of domains of contrasting densities. The result of the power spectrum indicates two interfaces: a shallow interface at an average depth of 3 km and deeper interface at an average depth of 8 km (Fig. 7). The interface at 8 km possibly indicates the mean depth of some high-density units such as rock units of schist belts.



Radially averaged power spectrum of Gravity anomaly

#### Conclusion from the maps

- From the 3D depth slice plot, Regional Bouguer anomaly plot and Radially Averaged power spectrum plot, it is inferred that the gold and iron mineralization present in the Chitradurga schist belt are of deeper origin more than 3km.

## SOURCE CODE AND RESOURCES

GitHub is used to upload Source Code.

GitHub source code can be accessed from the following link:.

[https://github.com/saudaminisahoo/HACKATHON\\_2025/blob/main](https://github.com/saudaminisahoo/HACKATHON_2025/blob/main)

## REFERENCES

Bhattacharya, B.K. (1967). Some general properties of potential fields in space and frequency domain: A review. *Geoexploration*, 5, 127-143.

Blakely, R. J. (1996). Potential theory in gravity and magnetic applications. Cambridge University Press.

Geological Survey of India (GSI), Standard Operating Procedures (SOP). (2011). A manual on SOP in mineral exploration (Mission II- A natural non- energy resource assessment): [www.portal.gsi.gov.in/gsiDoc/pub/draft\\_sop\\_minex.pdf](http://www.portal.gsi.gov.in/gsiDoc/pub/draft_sop_minex.pdf), Accessed in May, 2013.

Goodway, B. (2012). Introduction to this special section: Mining geophysics. *The Leading Edge*, 31(3), 288–290.

Jacobsen, B.H. (1987). A case for upward continuation as a standard separation filter for potential field maps. *Geophysics*, 52 (8) 1138– 1148.

Kesler, S.E. (2007). Mineral supply and demand into the 21st century. In proceedings for a work-shop on deposit modeling, mineral resource assessment, and their role in sustainable development. US Geological Survey circular Vol. 1294, pp. 55–62.

Lyngsie S. B, Thybo H. Rasmussen T. M. (2006). Regional geological and tectonic structures of the North Sea area from potential field modelling. *Tectonophysics*, 413, 147–170.

Maus, S. and Dimri, V. (1996). Depth estimation from scaling power spectrum of potential fields. *Geophys. J. Int.*, 124, 113-120.

Ng, Andrew. Neural Networks and Deep Learning (DeepLearning.AI).

Roy, A. and Arora, A. (2023). Predicting density in post-rift lithologic sequence from well-log data using multilayer perceptron neural network. IEEE, Second international conference on informatics.

Schmitt, D.R., Mwenifumbo, C.J., Pflug, K.A., & Meglis, I.L. (2003). Geophysical logging for elastic properties in hard rock: A tutorial. Hardrock seismic exploration. Edited by DW.

Singh, A., Mishra, P. K., & Sharma, S. P. (2019). 2D cooperative inversion of direct current resistivity and gravity data: A case study of uranium bearing target rock. *Geophysical Prospecting*, 67(3), 696–708.

Spector, A. and Grant, F.S. (1970). Statistical models for interpreting aeromagnetic data. *Geophysics*, 35, 293-302.

Wang, J., Meng, X., & Li, F. (2015). A computationally efficient scheme for the inversion of large scale potential field data: application to synthetic and real data. *Computers & Geosciences*, 85, 102–111.