

Bharathidasan University Tiruchirappalli – 620 023, Tamil Nadu

M.Tech. Remote Sensing and GIS

Course : 24MTRS-05 GEOGRAPHIC INFORMATION SYSTEMS (GIS)

Unit-5 Data Classification & Advanced GIS

Dr. K. Palanivel

Professor, Department of Remote Sensing

24MTRS-05: GEOGRAPHIC INFORMATION SYSTEM

UNIT I Basics of GIS - Definition - Evolution of GIS - Geospatial, spatial and non-spatial data - Components of GIS -Computer Hardware, Software Modules and Organisational Context of GIS - Data Structure in GIS - Types of Data (Points, Lines and Polygons) - Data sources - Ground and remote sensing survey - Data Structures (Raster data structures and Vector data Structures) – Database Structures.. 12 Hrs.

UNIT II GIS Data Input, Verification, Storage and Output - Spatial Data Input Processes and Devices (Sources of data, - Different Types of Data Entry methods, viz., Manual input, Run length code, Digitization, Automated Scanning, etc. – Vector to Raster conversion - Raster to Vector conversion - Input devices) - Entry of non-spatial data – Linking of Spatial & Non-spatial data – Element generation - Geodatabase creation - Data Verification (Errors of different types) - Correction (Rubber Sheet Transformation, Bilinear interpolation, Cubic Convolution, etc.) - GIS capabilities for Data correction – Data output (Types of Output, GIS Capabilities for output, Output devices) - File formats for raster and vector. 14 Hrs.

UNIT III Raster and vector data Interpolation and Analyses - Raster data analysis: Local, neighborhood and regional operations - Map algebra - Vector data analysis: Topological analysis, point-in-polygon, line-in-polygon, polygon-in-polygon - Proximity analysis: Buffering, Thiessen polygon - Non topological analysis: Attribute data retrieval and analysis – Concepts of SQL – ODBC - Basic Principles of Interpolation – Methods of Interpolation - Global Methods of Interpolation, Local Interpolation (Trend Surface Analysis) – Local Interpolation (Splines) - Optimal Interpolation (Kriging). 12 Hrs.

UNIT IV Surface Analysis, 3D Visualization and Network Modelling - Need for Three Dimensional Models - Methods of DEM & TIN - Products of DTM (Contour Maps, Shaded Relief Map, Maps Related to Slopes, Line Sight Maps, Drainage Analysis, Volume Estimation etc.) Usefulness of DEM/DTM. Overlay analysis, Capabilities (Point Operations, Regional Operations, Neighbourhood Operations) - Networking and Dynamic Segmentation – Applications, Minimum Distance Model, Maximum Covering Model (P-median model), Urban Transportation Planning Model. 14 Hrs.

UNIT V Data Classification & Advanced GIS - Principles – Types of Classification (Exogenous, Arbitrary, Idiographic & Serial) – Multivariate Analysis. Artificial Intelligence - Expert Systems - Object Oriented GIS - Web based GIS: Definition, merits - Architecture - Map server - Spatial data infrastructure - Spatial data standards - Free and open source - Proprietary GIS software. 12 Hrs.

UNIT VI Current Contours – GIS - Case studies (Not for Final Exam only for Discussion) - Recent scientific and technological development, advancement, Industrial application and Job opportunities.

REFERENCES

- 1. Burrough, P.A 1986: Principles of Geographical information Systems for Land Resources Assessment, Clarandone Press, Oxford.
- 2. Avery, T.V, Interpretation of Aerial Photography Burgass, Publishing Company.
- 3. Gautham, N.C 1970: Urban Landuse Study Through Aerial Photo binterpretations Techniques, Pink Publishing House, Mathura.
- 4. American Society of Photogrammetry, 1983: Manual of Remote Sensing (2nd Edition), ASP Falls Church, Virginia.
- 5. Campbell, J 1984: introductory Cartography, Printers Hall Englewood Cliffs, N.J
- 6. Dent B.D 1985: Principles of Thematic Map Design, Addition Wesley, Reading, Mass.
- 7. Freeman, H and GG.Pieroni 1980: Map Data Processing, Academic Press, New York.
- 8. Monmonier, M.A 1982: Computer Assisted Cartography Principles and Prospects, Prentice Hall, Englewood Cliffs, NJ
- 9. Tomlinson, RF Calkins, HS and D.F.Marble 1976: Computer Handling of Geographic Data, UNESCO, Geneva.
- 10. Graeme F. & Bonham Carter; Geographic information Systems for Geoscientists; Modelling with GIS, Pergamon.
- 11. Lo, C.P. and Yeung, Albert K.W., "Concepts and Techniques of Geographic Information Systems", Pearson, 2016.

OBJECTIVE

To understand the fundamentals, data handling, analysis techniques, and advanced applications of Geographic Information Systems (GIS).

COURSE OUTCOMES

- The concepts of GIS, including data types, components, and data structures.
- Develop proficiency in GIS data input, verification, storage, and output processes.
- Gain knowledge in raster and vector data analysis, including interpolation methods and proximity analysis.
- Understand three-dimensional modeling, surface analysis, and network modeling techniques.
- Explore classification methods, AI integration, web-based GIS, and the use of open-source and proprietary GIS software.

Unit-V

UNIT V Data Classification & Advanced GIS -Principles – Types of Classification (Exogenous, Arbitrary, Idiographic & Serial) – Multivariate Analysis. Artificial Intelligence - Expert Systems - Object Oriented GIS - Web based GIS: Definition, merits - Architecture - Map server -Spatial data infrastructure - Spatial data standards - Free and open source - Proprietary GIS software. 12 Hrs.

Importance & Principles of Data Classification

- For human understanding, classification is essential.
- Without classification or generalization, our brains becomes swamped by so much of details.
- Very often, literally we can not see the wood for the trees.
- Human brain seems to be limited in the number of classes that it can cope with Seven classes plus or minus two per level in the hierarchy – to be comfortable.
- Binary classification considered to be insensitve and of little general value.

Classification Methods

Classifying the ranges of values of a single continuous variables.

The different types of making class intervals are:

- Exogenous class intervals according to threshold values that are relevant to, but not derived from the data set under study.
 - Fail to resolve the variations in any specific area.
 - If exogenous intervals are derived from qualitative class definitions, they may not be unique and may in fact overlap.

Arbitrary class intervals

- chosen without any clear aim in mind
- very often the intervals are irregularly spaced and
- Chosen without a proper examination of the data.

Idiographic class intervals

- chosen with respect to specific aspects of the data set
- They include methods that attempt to divide the data into multi-modal groupings
- Using 'Natural Breaks', their location,
- Objective comparison is difficult not exactly followed from dataset to dataset.

Serial class intervals

- have limits that are direct mathematical relation to one another.
- intervals are chosen independent of the characteristics of the particular dataset,
- then the results from different samples or datasets will be strictly comparable.
 - Normal percentiles class intervals subdivide a normal distribution with mean μ and standard deviation, with classes centred on the mean.
 - Equal arithmetic intervals with no variation in class width, such as are normally used for elevation contours.
 - Equal intervals on reciprocal, trignometric, geometric, or logarithmic scales used when original data are not normally or rectangularly distributed.

Multivariate Analysis / Classification in GIS

- Statistical method for evaluating the relationship between two or more variables or classes of data samples from spatial data (satellite images).
- To establish the relationships within and between the classes
- Statistics are stored in a signature file.
- Two types of classification:

1) supervised and 2) unsupervised.

- Delineate training samples as classes supervised
- Multivariate statistics that are calculated on the input raster will identify clusters as classes unsupervised.

Transforming multivariate data: Principal component analysis (PCA)

- Principal component analysis (PCA) transforms the data attributes in a multiband raster from the input multivariate attribute space to a new multivariate attribute space whose axes are rotated with respect to the original space.
- The axes in the new space are uncorrelated i.e., segmented to independent groups with descending importance from the 1st to the last.
- Compresses the data by eliminating redundancy, emphasizes the variance within the bands of a raster, and makes the data more interpretable.
- The result of a PCA is a multiband raster with the same number of bands as the original raster (one band per axis in the new multivariate space).
- The first principal component will have the greatest variance, the second will show the second most variance not described by the first, and so forth.
- The **first three to four layers** of the resulting raster from the principal component function may describe more than **95%** of the variance with accuracy.
- The remaining layers of the principal component raster can be dropped.

Artificial Intelligence

Inducting the elements of intelligence to computer is called **artificial intelligence**.

Expert Systems

The computer which does intelligent job is called **expert** systems.

Artificial intelligence is designing intelligent computer systems; that is, systems that exhibit the characteristics, we associate with intelligence in human behavior **Definition:** "Artificial intelligence is the study of how to make computers do things at which, at the moment, people are better". *-Elian Rich*

Two basic ideas of AI are:

- 1. It involves studying the **thought processes of humans** (to understand what intelligence is);
- 2. It deals with **responding those processes via machines** (computer, robots, etc.).

AI is basically a theory of how the human mind works. Objectives of AI are:

- 1) Make machines smarter
- 2) Understand what intelligence is &
- 3) Make machines more useful.

Expert Systems for Geographic Information Systems

Major problem domains of ES-related efforts for GIS are:

- Map design,
- Terrain/feature extraction,
- Geographic / Spatial database management and
- Geographic / Spatial decision support systems.

Expert Systems (ES) are computer systems that advise on or help solve real-world problems which would normally require a human expert's interpretation. Such systems work through problems using a computer model of expert human reasoning.

Applications of ES:

- 1. Geographic data input used to :
 - extract features from imagery,
 - exploit the potential of automatic scanning of manuscript maps,
 - manage the editing of geographic data at the same time data are being captured, and
 - assess the quality of data being entered into the system.
- 2. Exploit knowledge about a user query and the GIS itself in order to speed the search through very large spatial databases.
- Development of intelligent user interfaces will make GIS responsive to user needs because the user will no longer have to become an expert in the use of GIS in addition to their own field of specialization.
- 4. Cartographic expert systems could be used to produce high quality maps which would effectively communicate information derived from a GIS.
- 5. Play a significant role in improving the efficiency and effectiveness of GIS.

Characterization of ES:

Expert systems should :

- interact with humans in natural language,
- function despite some errors in the data and uncertain judgemental rules,
- contemplate/consider multiple competing hypotheses simultaneously,
- explain their reasons for requesting additional information when needed &
- include any ability to learn using powerful learning ability machine learning.

Organization of ES:

- A knowledge-based system would separate domain-specific rules from the procedural language used for controlling program execution.
- This organization makes it much easier to encode and maintain facts and rules.
- In fact, PROLOG, a commonly used language for building expert systems, has been described as a database programming language.

Knowledge Representation:

At least in *three* ways:

- 1. Use the **formulas** of first-order predicate logic used to represent both *facts* and *rules*. Procedural declarative knowledge programming language PROLOG is an example
- 2. Use **semantic nets** (or semantic networks).- a means of modeling human associative memory objects are represented by nodes in a graph and the relations among them by labeled arcs.
- 3. Use **Frames** as data structures in which all knowledge about a particular object or event is stored together. Because the organization of knowledge is more modular, its accessibility is increased.

Use a **hybrid frame-based**, **semantic net method** of knowledge representation (KBGIS – Knowledge Base GIS).

Knowledge Exploitation:

- Solving steps are necessary for reaching a solution
- necessary to search through a "state-space" containing a large number of alternatives where each might lead to a solution.
- **Pattern matching** pattern rule and collection of objects.
- State-spaces can be thought of as being represented by a graph.
- Walking through the graph is analogous to searching for a path to reach a state that corresponds to a solution.
- They can also be searched in a **backward direction** by starting with a goal, or solution, state and then finding a path to the initial state.
- The **appropriateness of the approach** depends on the particular problem being addressed and the nature of the state-space.
- Two fundamental approaches are **forward chaining** and **backward chaining**.

Forward chaining is a data-driven approach.

- One of the major problems with implementing this approach is in specifying how to determine whether or not a fact is useful for the problem at hand.
- Backward chaining is goal-driven.
- The advantage of this kind of search is that the goal is known from the start and can be used to advantage when searching large statespaces.
- On the other hand, this introduces additional complexity into required pattern matching expressions.

Ref.: Vincent B. Robinson & Andrew U. Frank, 1987. Expert Systems for Geographic Information Systems. American Society for Photogrammetry and Remote Sensing, Photogrammetric Engineering and Remote Sensing, 0099-1112/87/5310-1435\$02.25/0, Vol. 53, No. 10, October 1987, pp. 1435-1441.

Live Maps

The map which can display the current situation / properties / latest changes, by involving recent data collected through a particular data collection network

For example, Hourly satellite data, hourly groundwater level data, continuous seismic data, cloud-pattern every 30 minutes, etc., received from such network continuously, preprocessed, interpolated for the entire study area, if needed, and then analysed to obtain a particular Action Plan map.

This map can:

- provide timely answers to the queries raised by the users &
- perform live interaction in order to get up-to-date information.

Virtues & Trends of GIS

- GIS has got a large number of vistas and advanced virtues
- Apart from many, GIS can perform a lot of sophisticated spatial operations.
- Integrate data from various sources RS, GPS, CAD, AM/FM, etc.
- Provide support for making decisions with a set of procedures in a problem solving environment.

CUSTOMIZATION IN GIS (Advancement : 1)

- Consist of easy user interface (GUI)
- User friendly environment
- Reduces multiple time-taking steps
- Hiding operations (unnecessary for the user)
- Easy to modify &
- No need to train the user rigorously.

CUSTOMIZATIONcontd...

- Generation of own interface Based on the nature of our project work – self designed
- Series of multiple tasks can be done easily as a simple task through a single mouse– click event
- Continuous and repeated processing till it attains a good result.

Process Automation (Advancement : 2)

Automation of certain regular processes to generate / extract data and derive products.

- Lineament extraction
- Drainage extraction
- Lineament / drainage classification
- Automated digitization,
- Automated density mapping of Lineament Intersections,
- Automated delineation of Watersheds, etc.

Decision Support Systems (Advancement : 3)

- DSS with multiple models
- Quick and easy access of multiple datasets as maps
- Complex querying and data retrieval
- Spatial display of vast non-spatial data quickly
- User defined options
- Easy access by non-GIS users too
- User Query based Information Retrieval System (QUBIS).

SPATIAL INFORMATION RETRIEVAL SYSTEM (Advancement : 4)

- Customized GIS package
- Assists the user to retrieve the spatial data easily and quickly
- User oriented built-in queries for easy retrieval
- Displays attribute data spatially by creating a link (Relate).

SPATIAL DECISION SUPPORT SYSTEM (SDSS)

- An interactive computer based system,
- designed to support the user
- in achieving a higher effectiveness of decision making
- while solving a semi structured spatial decision problem.

USER SIDE INPUTS

- User need to help the system to locate the data sets
- Select the model / ready made methodology
- Give the conditions, if necessary, then

SDSS does the job of

- Self extrapolating the data sets
- Generating the databases
- Self analyzing the databases,
- Generating action plans maps, tables, reports, etc.

Generic Characteristics / Usefulness of SDSS / DSS

- An explicit design to solve ill-structured problems
- Powerful and easy-to-use user interface
- Ability to continue analytical models flexibly with data
- Ability to explore the solution space by building alternatives
- Capability of supporting a variety of decision making styles and
- Allowing interactive and recursive problem solving.

DISTINGUISHING CAPABILITIES AND FUNCTIONS OF SDSS

- Provide mechanisms for the input of spatial data / updating of data
- Allow representation of the spatial relations and structures
- Include analytical techniques of spatial analysis, and
- Provide output in a variety of spatial forms including maps.

💐 Pulldown			
BASE MAPS THEMATIC MAPS	RESOURCES AND HAZARDS MAPS	RESOURCES-HAZARDS CONSERVATION CLEAR QUIT	
Form ROAD METAL	MINERAL RESOURCES SURFACE WATER RESOURCES GROUND WATER RESOURCES GEOTHERMAL PROVINCES SOIL EROSION LANDSLIDE SEISMIC CORRIDORS		
LIME STONE CLAY SILICA	WATERSHED WISE VILLAGE WISE	Select a Map for Wrapping Watershed Village Map Drainage	



Menus and Forms of WEGHARIS i.e. Western Ghats Resources Information System developed using **AML** (Arc Macro Language of PC ARC/INFO)

н.	с.	CI	мII	Π.
ь.	L	c,		ω.
	-	-		

ROAD METALS



WEGHARIS - Western Ghats Resources Information System developed using **AML** (Arc Macro Language of PC ARC/INFO)





74*40'34.16"E 13*30'20.06"N



NRDMS-QUBIS – Welcome page / Front Page

Advanced options are made available in NRDMS-QUBIS:

- 1. Easy & quick retrieval with GUI based mouse cursor events
- 2. Various Display Levels
- 3. Layer Wrapping
- 4. User Guidance for information display
- 5. Easy updation of data
- 6. User required Outputs-S/c&H/c
- 7. Spatial Display of non-spatial data.









OBJECT-ORIENTED GIS

Object-oriented data model is one application of object-oriented technology.

- It uses objects to organize spatial data.
- Unlike a geometric object of a point, line or area, an *OBJECT* is defined here as something that has a set of properties and can perform operations upon requests.
- That means, almost everything one uses in a GIS is an object.
- For e.g., a land use map is an object, which has properties, such as its coordinate system and feature type and can respond to requests such as zoom in, zoom out and a query.

Structural Aspects

Principles: To group objects : association, aggregation, generalization, instantiation and specialization aspects are required.

Association: Describes the relationships between objects of two types.

- If owner and land parcel represent two types of objects, the relationships between them can follow the rules that
- an owner can own one or more parcels and
- a parcel can be owned by one or more owners.

Aggregation: Asymmetric association in a whole-part relationship.

For e.g., block groups are connected to form a census tract and census tracts are connected to form a county.

Generalization: Identifies the commonality among objects, and groups objects of similar types into a higher-order type.

- For e.g., parcel, zoning and census tract maps may be grouped into a higher-order class called boundary.
- Grouping of objects forms a hierarchical structure, which organizes objects into classes and classes into superclasses and subclasses.

Instantiation: An object of a class can be created from an object of another class.

For e.g., a high-density residential area object may be created from a residential area object.

Specialization: Differentiates objects of a given class by a set of rules.

For e.g., roads may be separated by average daily traffic volume.

Behavioural aspects of objects:

Inheritance: It is the basic principle in explaining the behaviours of objects:

- subclasses inherit properties and operations from a superclass, and
- objects inherit properties and operations from a subclass.
- For e.g., the residential area is a superclass and low-density area and high-density area are the subclasses.
- All properties of the class residential area are inherited by its own subclasses.
- Through inheritance, properties need only be defined once in the class hierarchy.

Encapsulation: Encapsulation refers to the mechanism to hide the properties and operations of an object so that the object can perform an operation by responding to a predefined message or request.

For e.g., a polygon object can respond to a request called ReturnCenter (return the center of) by returning the physical center of the polygon.

Polymorphism: It allows the same operation to be implemented in different ways in different objects.

For e.g., the same request called GetDimension (get the dimension of) can be sent to a point, a line, or a polygon but the result differs depending on the feature type.

The number 0 is returned if the object is point,

1 if the object is a line, and

2 if the object is a polygon.

Web based GIS

- An advanced form of geospatial information system
- A combination of **software** and **hardware**
- These enables the distribution of maps, attribute data and geographic processing capabilities
- over a network and the internet using standard web communication protocols (http).
- Information exchange takes place between a server and a client,
- The server possesses a unique Uniform Resource Locator (URL) that clients use to find it on the web.

Ref. Link: https://www.nrsc.gov.in/sites/default/files/pdf/ebooks/QRG_on_WebGIS.pdf

Merits of Web GIS:

- Web GIS broadens the reach of GIS by making its features available to a larger audience.
- Web GIS is not intended to outperform or compete with desktop GIS but to **extend GIS capabilities to everyone**.
- Web GIS reduces the necessity for users to install complex software since most tasks can be carried out within web browsers or mobile applications.
- A web service is a piece of code operating on the server that executes actions in response to client requests.
- With the Internet's help, clients can access geospatial information online, regardless of the distance between the server and client.
- Web GIS offers global sharing of Geographic data and maps.

The key elements of Web GIS include:

- **1. Server Accessibility:** The server has a URL that allows clients to locate it on the web.
- 2. Client Requests: Clients send requests to the server according to HTTP specifications.
- **3.** Server Response: The server performs the requested GIS operations and responds to the client via HTTP.
- 4. Response Formats: Responses can be provided in various formats, including HTML, binary images, XML (Extensible Markup Language), or JSON (JavaScript Object Notation).
- Web GIS is characterized by being object-oriented, distributed, and interoperable.
- It encapsulates GIS data and functions into objects that facilitate distributed computing, storing on multiple servers and exchanging and interacting using a consistent communication protocol.

Map server

- Map Server is an Open Source platform for publishing spatial data and interactive mapping applications to the web.
- Originally developed in the mid-1990's at the University of Minnesota, MapServer is released under an MIT-style license, and runs on all major platforms (Windows, Linux, Mac OS X).
- Map Server is not a full-featured GIS system, nor does it aspire to be.
- Map Server is a project of the Open Source Geospatial Foundation.
- Ref. Link : <u>https://mapserver.org/</u>

Spatial Data Infrastructure (SDI)

- A spatial data infrastructure (SDI), also called geospatial data infrastructure, is a data infrastructure implementing a framework of geographic data, metadata, users and tools that are interactively connected in order to use spatial data in an efficient and flexible way.
- SDIs are made up of a variety of components, including:
 - Data: Geographic data, metadata, and other digital resources
 - Technology: Hardware, software, and other technologies
 - **Policies:** Standards, institutional arrangements, and government policies
 - **People:** Users, organizations, and other human resources.
- SDIs can be used by a wide range of users, including: Government agencies, Commercial businesses, Non-profit organizations, Academic institutions, and Citizens.
- SDIs can be developed at various levels of scale and complexity, such as regional, state, agency, or national. Some examples of SDI components include geoportals, web GIS, and mobile GIS.
- Some challenges to SDIs include: Legal frameworks, Data protection, Copyright, Ownership, and Human interoperability.

Spatial data standards

Spatial data standards are rules and guidelines that help make it easier to share, transfer, and use spatial data:

- Spatial Data Transfer Standard (SDTS): A standard from the American National Standards Institute (ANSI) that specifies how to transfer spatial data between different computer systems. It includes information on the format, structure, and content of the data.
- Content Standard for Digital Geospatial Metadata (CSDGM): A standard used by NSDI stakeholders to format metadata records.
- **ISO 19131:** A standard that sets out the purpose of a dataset.
- OGC GeoPackage 1.0: A standard that defines GeoPackages for exchanging data. GeoPackages can be used to access and update data without translating it into an intermediate format.
- **Geographic Information Framework Data Standard:** A standard with parts that cover cadastral, digital orthoimagery, elevation, and geodetic control.

Other spatial data standards include:

- **IS 13393**, which is a standard for representing geographic point locations by coordinates
- IS 16439, which is a metadata standard for geospatial information
- IS 16554, which is a data exchange standard for geospatial information
- IS 16626, which is a standard for Geographic Information - Geography Markup Language (GML)

NSDI Portal link:

https://www.nsdiclearinghouse.gov.in/erdas-apollo/nsdiportal/nsdistandards.html

Open Source GIS vs Proprietary GIS

- With open source software, users are free to utilize or modify the code as needed.
- Proprietary software comes with restrictions such as no access to modify the code and certain facilities are limited in versions to variety of users.
 - QGIS Open source GIS
 - GRAM ++ -developed by CSRE, IIT-B with the funding of DST, Government of India
 - ArcGIS Pro (ESRI)
 - Geomedia (Hexagon Geospatial)
 - MapInfo Professional (Precisely)
 - Global Mapper (Blue Marble).

Thankyou

For your kind cooperation Patient listening & Learning GIS

Let us enjoy using GIS more efficiently & effectively...