

DATA ANALYSIS MODELING

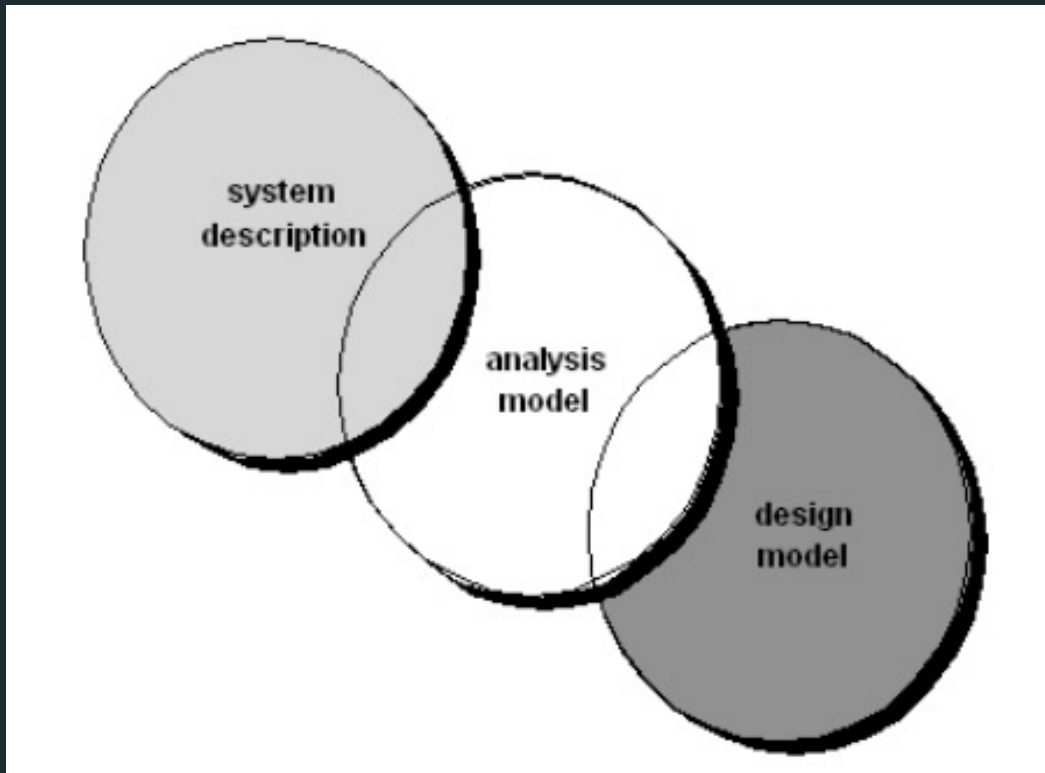


Thip Rattनाविलay

DSC500 - Introduction to data science

REQUIREMENTS ANALYSIS

- Requirements analysis
 - ✓ Specifies software's operational characteristics
 - ✓ Indicates software's interface with other system
 - ✓ Elements establishes constraints that software must meet
-
- Requirements analysis allows the software engineer (called an analyst or modeler in this role) to:
 - ✓ elaborate on basic requirements established during earlier requirement engineering tasks
 - ✓ build models that depict user scenarios, functional activities, problem classes and their relationships, system and class behavior, and the flow of data as it is transformed.



(Redirect Notice)

ANALYSIS MODEL

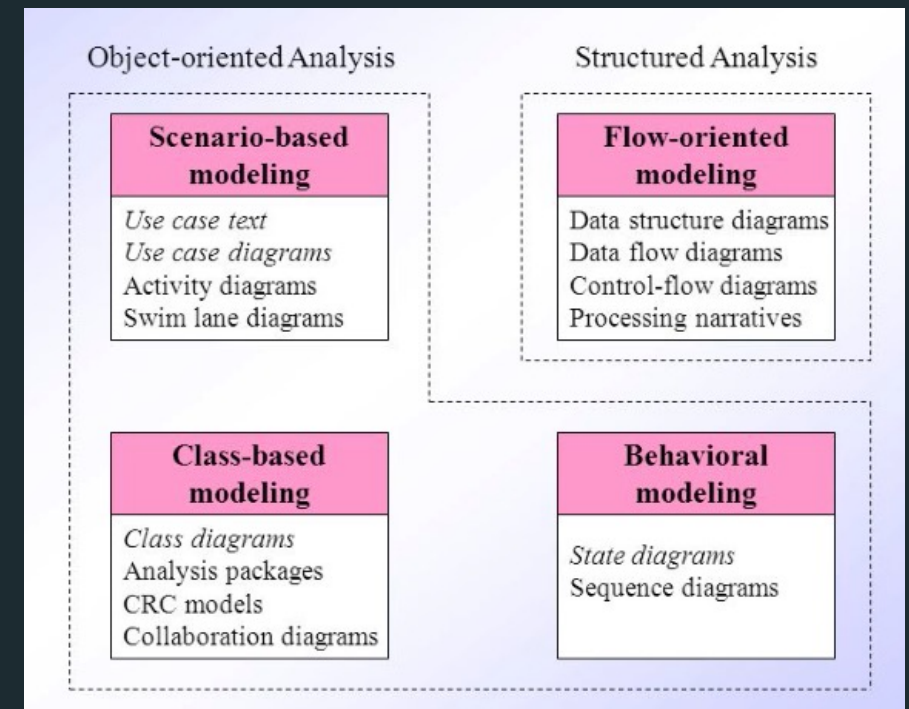
ANALYSIS MODEL: a representation of requirements in terms of text and diagrams depicting requirements for data, function and system behavior in a way easy to understand and review for correctness, completeness and consistency

RULES OF THUMB

The model should focus on requirements that are visible within the problem or business domain. The level of abstraction should be relatively high (focus on the "WHAT", not "HOW"). Each element of the analysis model should add to an overall understanding of software requirements and provide insight into the information domain, function and behavior of the system. Delay consideration of infrastructure and other non-functional models until design. Minimize coupling throughout the system. Be certain that the analysis model provides value to all stakeholders.

- Requirements analysis
- Flow-oriented modeling
- Scenario-based modeling
- Class-based modeling
- Behavioral modeling

MODELING TYPE



(Redirect Notice)

Logistic Regression Model

Logistic Regression is great for multiclass classification because Scikit-learn encodes the target labels automatically if they are strings (Smith, 2019).

```
from sklearn.preprocessing import scale
from sklearn.model_selection import train_test_split

X = X.drop(columns='id') #id is our index and won't help our model
X = scale(X)

X_train, X_test, y_train, y_test = train_test_split(
    X, y, train_size=0.75, test_size=0.25, random_state=42,
    shuffle=True)
```



OVERALL OBJECTIVES

THREE PRIMARY OBJECTIVES

- To describe what the customer requires
- To establish a basis for the creation of a software design
- To define a set of requirements that can be validated once the software is built
- All elements of an analysis model are directly traceable to parts of the design model, and some parts overlap

REFERENCES

- Smith, K. (2019, February 09). Predictive Modeling: Picking the best model. Retrieved November 19, 2020, from <https://towardsdatascience.com/predictive-modeling-picking-the-best-model-69ad407e1ee7>

(Smith, 2019)
- Redirect Notice. (n.d.). Retrieved November 19, 2020, from <https://www.google.com/url?sa=i>