

# ANL312 Text Mining and Applied Project Formulation

**Level:** 3

**Credit Units:** 5 Credit Units

**Language:** ENGLISH

**Presentation Pattern:** EVERY JULY

## Synopsis:

ANL312 Text Mining and Applied Project Formulation covers Text Mining as well as topics related to the formulation of the Applied Project such as Applied Project Research and report writing. The aim of this course is to broaden students' exposure to the breadth and depth of text mining and its applications and also to enable the students to carry out a prototype of the Applied Project so that they are more prepared for the Applied Project. Specific software will be used intensively in the course to provide hands-on applications of the topics covered.

## Topics:

- Applied Project Research
- Applied Project Report Writing
- Overview of Text Mining
- Data Preparation: Parsing and Feature Extraction
- Text Transformation: Term-by-Document matrix, Frequency weightings, term weightings, document filtering
- Bag of Words Approach
- Text Mining using Predictive Modeling
- Entity Extraction
- Document Classification
- Document Clustering
- Topic Extraction: singular value decomposition
- Sentiment analysis and Opinion mining

## Learning Outcome:

- Discuss text mining concepts and the various applications of text mining
- Evaluate the usefulness of several text specific transformation techniques (e.g., Zipf's Law and term-by-document matrix)
- Explain the process of document clustering, latent semantic indexing and topic extraction
- Compare the different approaches to text mining
- Apply clustering, singular value decomposition, latent semantic indexing and topic extraction
- Implement the various approaches to text mining using appropriate text mining software
- Propose an appropriate research topic for an analytics applied project
- Construct a prototype as part of an applied project development
- Draft and submit a report on the prototype of the Applied Project

## Assessment Strategies (Daytime Class):

Components	Description	Weightage Allocation (%)
------------	-------------	--------------------------

Overall Continuous Assessment	PRE-COURSE QUIZ 1	2
	PRE-CLASS QUIZ 1	2
	PRE-CLASS QUIZ 2	2
	PARTICIPATION 1	6
	TUTOR-MARKED ASSIGNMENT 1	18
	GROUP BASED ASSIGNMENT 1	20
Overall Examinable Components	ECA	50
<b>Total</b>		<b>100</b>

**Assessment Strategies (Online Class):**

Components	Description	Weightage Allocation (%)
Overall Continuous Assessment	PRE-CLASS QUIZ 1	2
	PRE-CLASS QUIZ 2	2
	PRE-COURSE QUIZ 1	2
	DISCUSSION BOARD 1	10
	GROUP BASED ASSIGNMENT 1	10
	PARTICIPATION 1	6
	TUTOR-MARKED ASSIGNMENT 1	18
Overall Examinable Components	ECA	50
<b>Total</b>		<b>100</b>