



Web Analytics and Natural Language Processing (WANLP)

Description détaillée du contenu du module WANLP :

Présentation des cours, des TD et des TP : Notions & Concepts | Principes et Techniques | Langages, Environnements & Librairies | Projet à réaliser le cadre de ce module

Professeur, BOURAMOUL Abdelkrim

Département IFA, Faculté NTIC

abdelkrim.bouramoul@univ-constantine2.dz

www.bouramoul.com

Etudiants concernés

Faculté/Institut	Département	Niveau	Spécialité
NTIC	IFA	Master 1	SDIA

Syllabus du module WANLP

Intitulé du Master

**Science de Données et
Intelligence Artificielle**

Semestre

S2

Unité d'Enseignement

UEF2

Intitulé de la Matière

**Web Analytics and Natural
Language Processing**

Code

WANLP

Crédit

6

Coefficient

3

Nombre d'heures d'enseignement

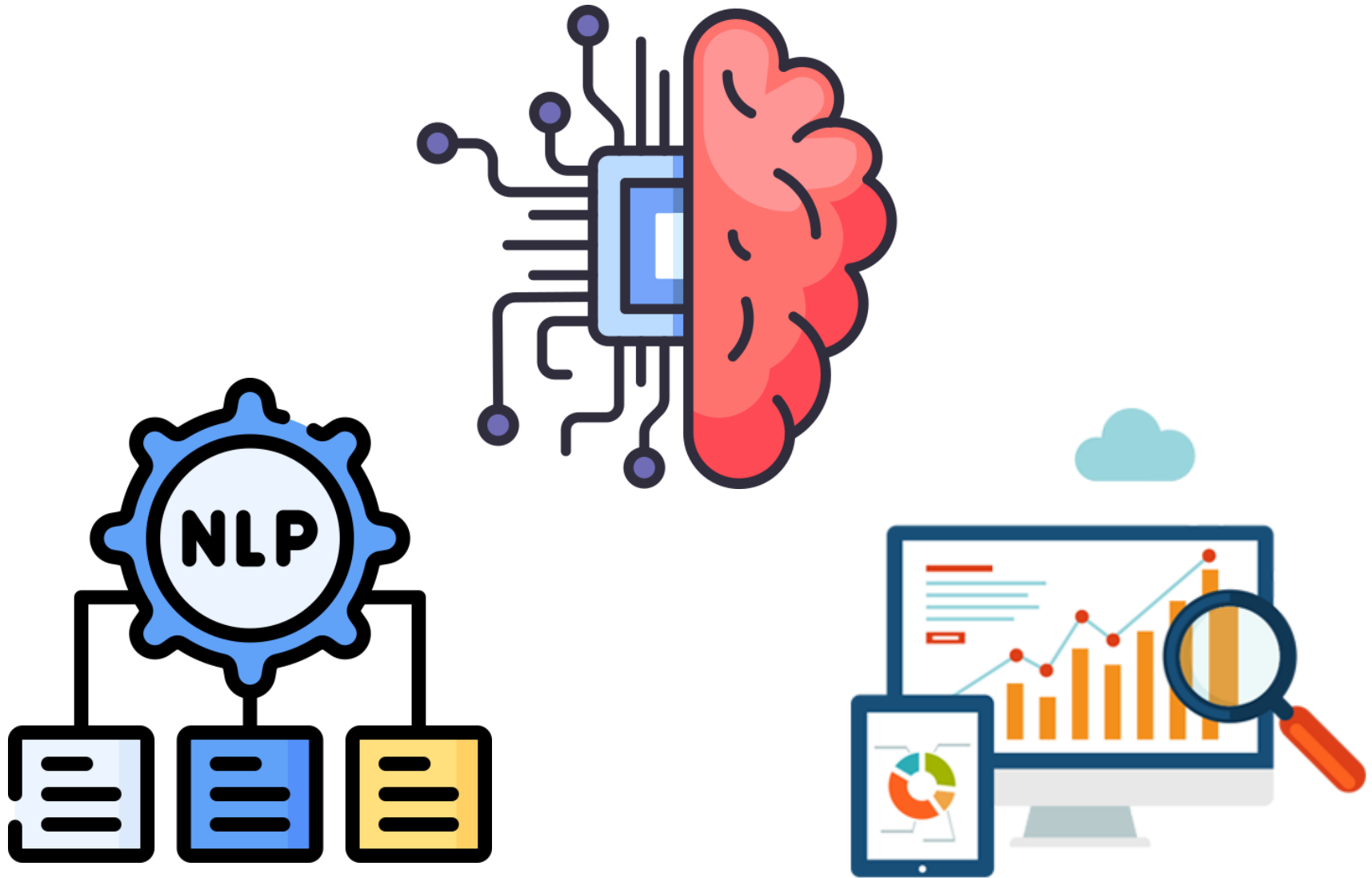
- **63 H sur 12 semaines**
- **Cours : 1H30/ semaine**
- **TD: 1H30/semaine**
- **TP: 1H30/semaine**

Objectifs de l'enseignement

- Comprendre les bases de l'IA, du TALN et du Web Analytics
- Maîtriser les techniques avancées en TALN et en Analyse Web.
- Développer des compétences pratiques en programmation Python et en utilisation des bibliothèques courantes.
- Intégrer ces connaissances dans des projets pratiques combinant le TALN, l'Analyse Web et les modèles génératifs.

Connaissances préalables recommandées

- **Théorie des langages de programmation et applications (THP)**
- **Probabilités et statistiques, Algèbre et analyse, Apprentissage automatique, Programmation (surtout en Python et Java)**



Chapitre 01

Introduction

Chapitre 01 : 02 Semaines

Cours : Introduction à l'IA, au NLP et au Web Analytics

- Concepts de base de l'Intelligence Artificielle, du Traitement Automatique du Langage Naturel et de l'Analyse Web.
- Importance et rôle de l'IA, du NLP et de l'analyse web dans les applications modernes.
- Vue d'ensemble des techniques et des applications dans chaque domaine.

TD : Applications de l'IA, du NLP et du Web Analytics dans la vie quotidienne

- Discussion en groupe sur les différentes applications, système et plateforme de l'IA, NLP et Web Analytics
- Exercice sous forme de quiz interactif pour évaluer la compréhension des concepts fondamentaux de l'IA, du NLP et de l'analyse web abordés dans le cours.

TP : Prise en Main des Outils

- Présentation des outils et des environnements de développement nécessaires pour le module (Python, Jupyter Notebook).
- Installation et configuration des bibliothèques Python essentielles pour le TALN et l'analyse web (NLTK, SpaCy, Pandas).



Chapitre 02 & 03

NLP

Chapitre 02 : 02 Semaines

Cours : Fondements et Concepts de base du NLP

- Concepts fondamentaux du Traitement Automatique du Langage
- Tokenisation, Normalisation et suppression du bruit dans le prétraitement de texte.
- Différentes techniques et approches utilisées dans le NLP pour traiter les données textuelles.

TD : Exercices sur le prétraitement de texte

- Exercices de mise en pratique les concepts abordés dans le cours.
- Exercices de tokenisation : diviser un texte en mots ou en phrases.
- Exercices de normalisation : uniformiser la casse, la ponctuation, etc.
- Exercices de suppression du bruit : éliminer les caractères indésirables, les mots vides, etc.

TP : Utilisation des bibliothèques Python pour le NLP (ex : NLTK)

- Fonctionnalités principales de la bibliothèque NLTK (Natural Language Toolkit) en Python.
- Utilisation de NLTK pour réaliser des tâches de base en NLP, telles que la tokenisation, le stemming et la lemmatisation.
- Application des techniques de prétraitement de texte apprises dans le TD en utilisant les fonctions et les modules de NLTK.

Chapitre 03 : 02 Semaines

Cours: Techniques Avancées en NLP : Modélisation du langage et classification de texte

- Techniques avancées en modélisation du langage naturel.
- Méthodes de vectorisation de texte pour représenter les mots et les phrases.
- Classification de texte et analyse de sentiment en utilisant des techniques d'apprentissage supervisé.

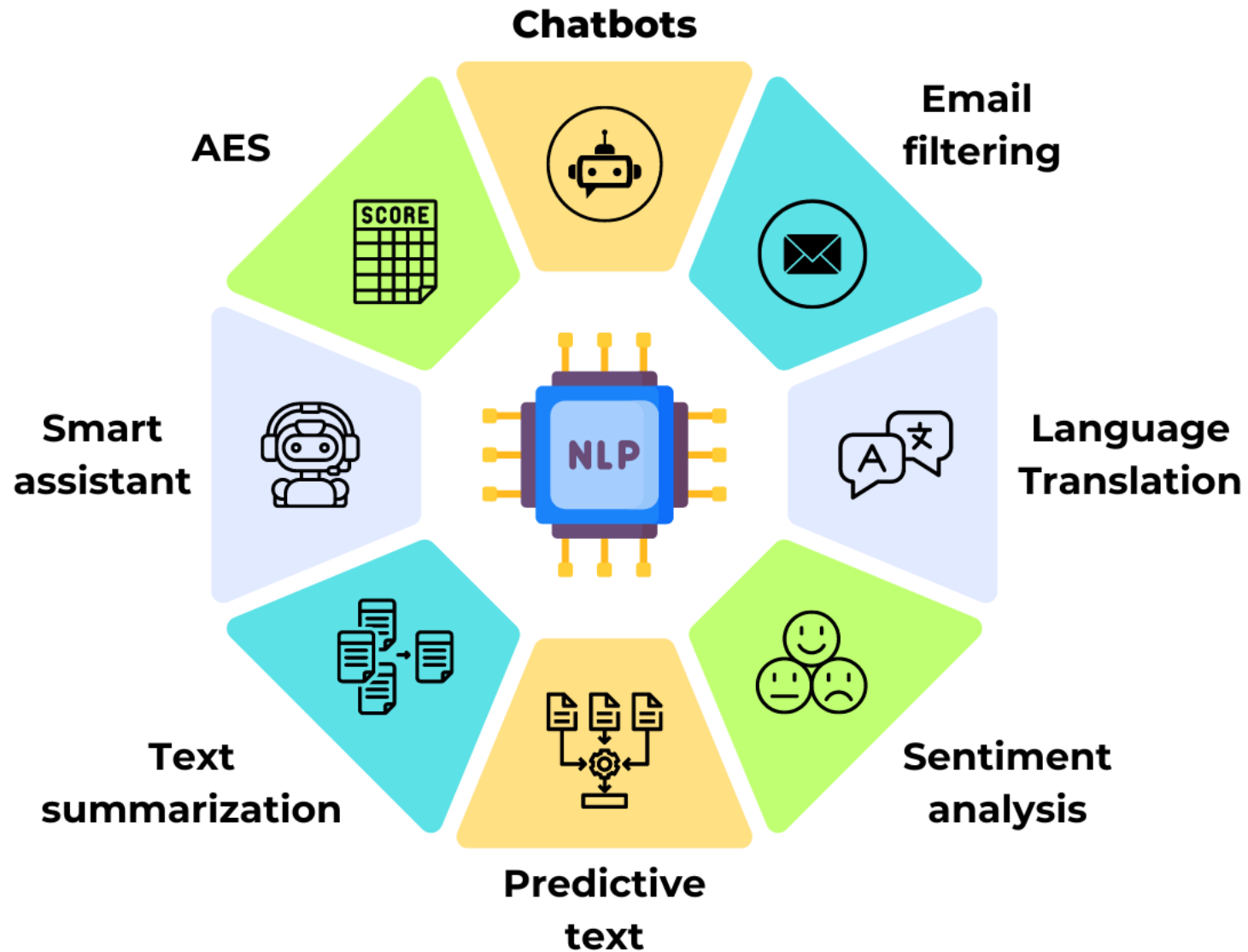
TD : Exercices pratiques sur la classification de texte et l'analyse de sentiment

- Mise en pratique les concepts abordés dans le cours.
- Classification de texte : entraînement d'un modèle pour classer des documents textuels dans des catégories prédéfinies.
- Analyse de sentiment : utilisation de techniques de NLP pour déterminer le sentiment (positif, négatif, neutre) dans des textes.

TP : Exercices sur l'implémentation d'un modèle de classification de texte simple

- Présentation des concepts de base de la classification de texte.
- Implémentation d'un modèle de classification de texte simple en utilisant des bibliothèques Python telles que Scikit-Learn.
- Techniques de classification de texte sur un problème concret, (Classification des critiques de films en positive ou négative).

Domaines d'application du NLP



Chapitre 04 & 05

Web Analytics

Chapitre 04 : 02 Semaines

Cours : Fondements et Concepts de base du Web Analytics

- Introduction aux principes fondamentaux de l'analyse web.
- Objectifs de l'analyse web et son importance dans les stratégies numériques.
- Différentes métriques utilisées en analyse web pour évaluer la performance d'un site web.

TD : Exercices sur les métriques clés en Web Analytics

- Comprendre comment calculer les métriques clés en analyse web.
- Calcul du taux de rebond, du taux de conversion, du temps moyen passé sur le site, etc.
- Interprétation des métriques et leur utilité dans l'optimisation des performances d'un site web.

TP : Exploration des données web avec Google Analytics

- Présentation de l'interface et des fonctionnalités de Google Analytics.
- Exploration des données web à l'aide de Google Analytics.
- Exploration des différentes sections de Google Analytics, examinent les rapports préétablis et créent des rapports personnalisés pour analyser les performances d'un site web.

Chapitre 05 : 02 Semaines

Cours: Techniques Avancées en Web Analytics

- Approfondissement des techniques avancées en analyse web.
- Segmentation avancée des utilisateurs pour cibler des audiences spécifiques.
- Analyse des entonnoirs de conversion pour comprendre le parcours des utilisateurs sur un site web et identifier les points de conversion.

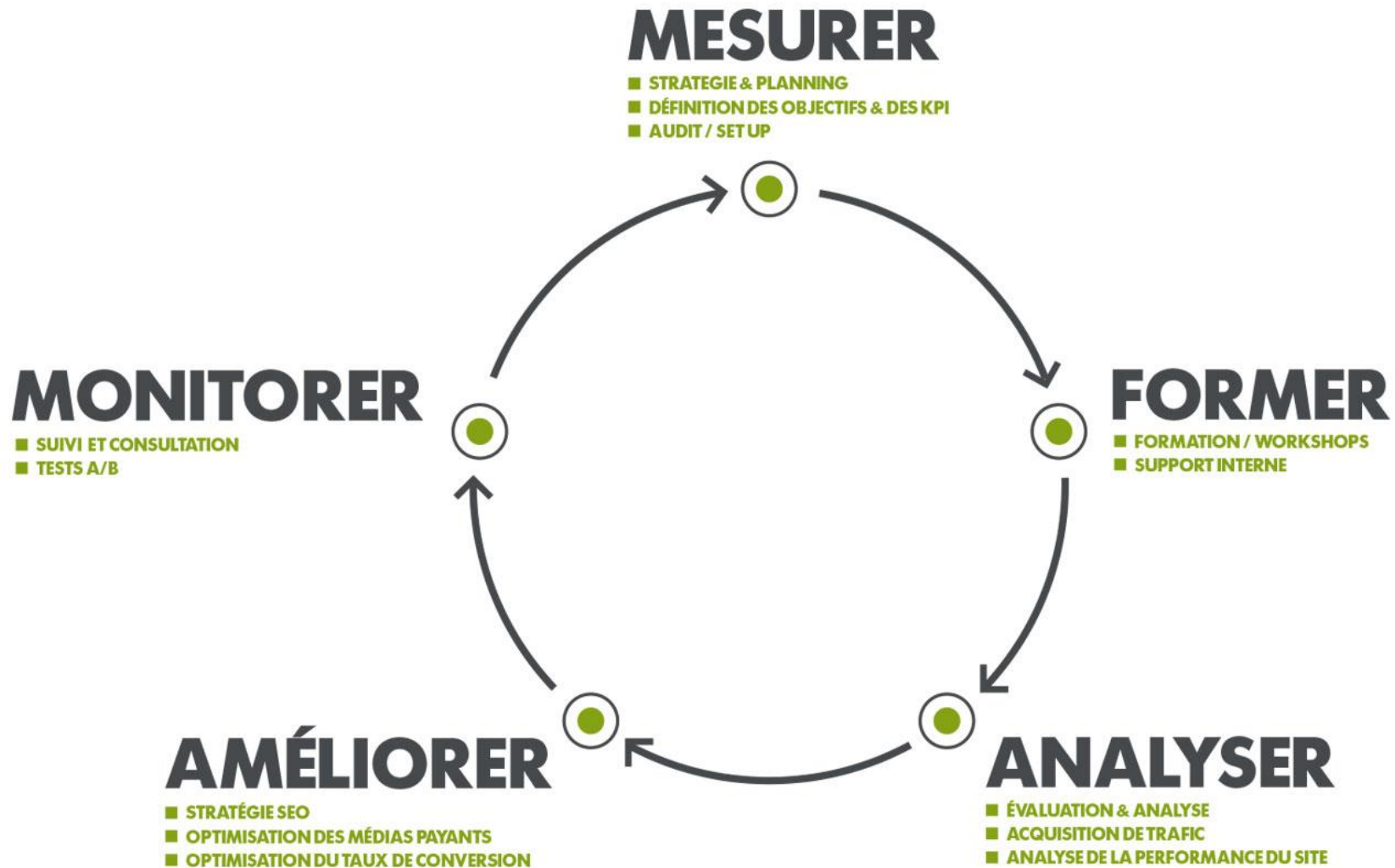
TD: Exercices sur l'analyse des entonnoirs de conversion

- Application des concepts abordés dans le cours.
- Création d'entonnoirs de conversion, analyse des taux de conversion à chaque étape de l'entonnoir, l'identification des points de sortie, etc.
- Interprétation des résultats et les actions à entreprendre pour améliorer les performances de conversion.

TP: Utilisation d'outils d'analyse web pour prédire le comportement des utilisateurs

- Outils d'analyse web avancés pour la prédiction du comportement des utilisateurs.
- Utilisation de ces outils pour prédire les actions futures des utilisateurs en fonction de leur comportement passé.
- Exploration des outils et réalisent des prédictions de comportement utilisateur en utilisant des données réelles ou simulées.

Démarche et objectifs du Web Analytics



Chapitre 06

IA Générative

Chapitre 06 : 02 Semaines

Cours: Modèles d'IA Génératifs et Intégration des Techniques NLP

- Principes fondamentaux des modèles génératifs en intelligence artificielle.
- Différentes architectures de modèles génératifs, (Réseaux Générateurs Adverses (GANs), Variational Autoencoders (VAEs), etc.)
- Applications des modèles génératifs dans divers domaines, tels que la génération d'images, de texte, de musique, etc.

TD : Exercices pratiques sur l'utilisation des GANs

- Comprendre le fonctionnement des Réseaux Générateurs Adverses (GANs).
- Génération d'images à partir de données bruitées, la génération de nouveaux échantillons à partir d'un ensemble de données existant, etc.
- Défis et les limites des GANs, et meilleures pratiques pour les utiliser efficacement.

TP : Création d'un modèle de langage génératif simple

- Concepts de base des modèles de langage génératifs.
- Création d'un modèle de langage génératif simple en utilisant une bibliothèque Python comme TensorFlow ou PyTorch.
- Implémentation et entraînement leur propre modèle de langage génératif sur un petit ensemble de données textuelles.

Chapitre 07

Projet de synthèse

Chapitre 07 : Tout le Semestre

Projet : Projet d'Intégration des Techniques de l'PNL et du Web Analytics en IA

- Introduction aux projets intégrés combinant le TALN, l'analyse web et l'IA génératifs.
- Présentation du cahier des charges : objectifs du projet & résultats attendus.
- Formation des équipes et attribution des rôles.
- Sessions de travail en groupe pour l'avancement des projets.
- Présentation et évaluation des projets réalisés





Université Constantine 2
جامعة قسنطينة 2

Fin

Description détaillée du contenu du module WANLP