

Projet Intégré 2BI : Mise en place d'un système d'information décisionnel

Encadré par:

Dr. Khouloud TALMOUDI (khouloud.talmoudi@esprit.tn)

Dr. Mohamed Aymen BEN HAJKACEM (medaymen.hajkacem@esprit.tn)

Description

- ▶ L'objectif de PI de cette année est vous former à mettre en place un système d'information décisionnel qui comporte deux parties, BI et analytique. La première partie consiste à créer un data warehouse et élaborer un tableau de bord interactif. La deuxième partie consiste à appliquer des techniques de data mining afin d'extraire des informations pertinentes et utiles à la prise de décision.
- ▶ Ce projet a pour objectif de donner aux décideurs un accès simple et rapide à la consultation des tableaux de bord. Il permet aussi de transformer les données stockées dans le data warehouse en connaissances utiles dans le processus décisionnel.
- ▶ Vous êtes invités dans ce projet à :
 - ❖ Traiter une problématique d'actualité via un jeu de données réel.
 - ❖ Utiliser **Talend** dans l'intégration des données et l'alimentation du data warehouse.
 - ❖ Visualiser les données en utilisant l'outil **Power BI**.
 - ❖ Implémenter les techniques de data mining avec **Python**.

Programme

- Organisation des équipes.
- Rédaction du cahier de charge
- Conception du data warehouse
- Alimentation du data warehouse avec Talend
- Visualisation du data warehouse avec Power BI
- Préparation des données
- Modélisation des données avec les techniques de Data mining

Etape 1 : Organisation des équipes

Durée : une semaine

Objectif : Organisation des équipes et choix des sujets.

- Formation des groupes de 5 étudiants.
- Choix d'un chef par groupe.
- Présentation des objectifs du PI.
- Chaque groupe doit fixer un sujet et trouver une source de données.

Etape 2 : Rédaction du cahier de charge

Durée: une semaine

Objectif : Chaque groupe doit préparer un cahier de charge qui comporte les éléments suivants.

- Le cadre général du sujet.
- Une étude de l'existant pour fixer la problématique du projet.
- Une description détaillée de la solution proposée.
 - ❖ L'analyse des besoins.
 - ❖ Les étapes du projet.

Livrable : Chaque groupe doit remettre un document pdf contenant le cahier de charge de son projet.

Etape 3 : Conception du data warehouse

Durée : une semaine

Objectif : Chaque groupe doit concevoir la structure du data warehouse.

- Définir les faits et les mesures.
- Définir les dimensions.
- Elaborer le schéma en étoile.
- Créer la structure du schéma dans un SGBD.

Livrable: Chaque groupe doit remettre un document pdf contenant le schéma du data warehouse.

Etape 4 : Alimentation du data warehouse avec Talend

Durée : 2 semaines

Objectif : Chaque groupe doit préparer les données en utilisant l'outil Talend.

- Télécharger les données sources.
- Configurer la connexion avec Talend
- Appliquer des transformations sur les données.
- Alimenter le data warehouse dans le SGBD.

Livrable : Chaque groupe doit remettre un document pdf contenant des captures d'écran qui illustrent les étapes de l'alimentation du dw.

Etape 5 : Visualisation du data warehouse avec Power BI

Durée : 2 semaines

Objectif : Chaque groupe doit créer les tableaux de bord avec l'outil PowerBI.

- Sélectionner les données à visualiser.
- Importer les données dans le Power BI.
- Créer un tableau de bord interactif à partir des données.
- Interpréter et analyser les graphiques.

Livrable : Chaque groupe doit remettre un document pdf contenant des chartes graphiques avec leurs interprétations.

Etape 6 : Préparation des données

Durée : 2 semaines

Objectif : Chaque groupe doit préparer et nettoyer les données avant d'appliquer les techniques de data mining.

- Sélectionner les données à partir du data warehouse
- Importer les données dans un outil de data mining.
- Appliquer les techniques de nettoyage pour traiter les données
- Appliquer les techniques de normalisation.
- Appliquer les techniques de réduction de la dimensionnalité.

Livrable : Chaque groupe doit remettre un document contenant des captures d'écran qui illustrent les étapes de la préparation des données.

Etape 7: Modélisation des données

Durée : 2 semaines

Objectif : Chaque groupe doit appliquer les techniques de data mining sur les données nettoyées.

- Choisir la tâche d'apprentissage et la technique utilisée.
- Appliquer la technique sur les données.
- Interpréter les résultats.
- Évaluer et comparer les résultats avec les techniques existantes.

Livrable : Chaque groupe doit remettre un document contenant des captures d'écran qui illustrent les étapes de la modélisation des données.

Evaluation

L'évaluation des projets se fera autour des points suivants :

- Rapport incluant une présentation de la problématique abordée, une description de la solution proposée qui intègre l, la préparation, la modélisation et la visualisation des données, ainsi qu'une présentation de la solution analytique.
- Soutenance finale de présentation du travail effectué.

Déroulement du projet

- ▶ Le chef de chaque groupe se chargera de la remise des travaux sur l'espace de PI sur Moodle.
- ▶ Les livrables de chaque étape et le rapport final doivent être remis sur l'espace dédié à cela sur Moodle en respectant les deadlines proposés.
- ▶ Le travail est effectué par groupe: tous les membres du groupe doivent impérativement s'impliquer dans la réalisation de toutes les étapes de projet.
- ▶ Tous les livrables sont évalués par l'enseignant encadrant.

Calendrier

- 02 Février 2022: Démarrage des projets (tous les étudiants doivent avoir un projet)
- Février ; Mars ; Avril(date à préciser): Soumission des livrables (rapports, code, interfaces...)
- Mai (date à préciser): Soutenances de fin de projet.

Bon travail 😊