

# INF8770

## Technologies multimédias

### A2025 - Travail pratique #1

#### Compression sans perte: texte

---

#### Remise du travail :

- Au plus tard, le 29 septembre 2025 à 09h sur Moodle - **aucun retard accepté.**

#### Documents à remettre :

- Votre code source ainsi qu'un rapport en **PDF** dans une archive (.zip/.7z/...). Un rapport sous un autre format se verra appliquer une pénalité. Le nombre maximale des pages accepté est de 10 pages sans annexe, l'annexe n'est pas pris en compte lors de la correction.

#### Autres directives :

- Il vous est interdit d'utiliser du code écrit par d'autres équipes.
- Il vous est permis d'utiliser du code trouvé sur internet et des bibliothèques externes, mais vous devez nous donner toutes les références.
- Le code lui-même n'est pas évalué, mais il doit être remis au complet.

#### Pénalités éventuelles :

- **Rapport :**
    - Rapport dépassant la limite : **-0,5 pt / page supplémentaire.**
    - Rapport non soumis sur Turnitin : **-0,5 pt.**
    - Rapport hors gabarit (police, interligne, etc.) : **-1 pt.**
    - Figure sans légende ou sans unité : **-0,5 pt / élément.**
    - Graphique ou tableau manquant : **-0,5 pt / élément.**
  - **Citations et intégrité :**
    - Références absentes ou non citées : **-0,5 pt.**
    - Code IA non cité : **note 0 sur la section concernée.**
    - Plagiat non cité / détecteur copié : **note finale 0.**
- 

#### Préambule :

Ce travail pratique s'inscrit dans le cadre du cours INF8770 — Technologies multimédias et a pour objectif de relier les notions fondamentales sur la structure des données multimédias à leur compression efficace, en mettant l'accent sur les méthodes étudiées aux chapitres 1 et 2.

Ce travail pratique inclut une évaluation des qualités 3 et 5 du Bureau canadien d'agrément des programmes de génie (BCAPG).

- **Qualité 2 Analyse des problèmes** : capacité d'utiliser les connaissances et les principes appropriés pour identifier, formuler, analyser et résoudre des problèmes d'ingénierie complexes et en arriver à des conclusions étayées.
- **Qualité 5 Utilisation d'outils d'ingénierie** : capacité de créer et de sélectionner des techniques, des ressources et des outils d'ingénierie modernes et de les appliquer, de les adapter et de les étendre à un éventail d'activités simples ou complexes, tout en comprenant les contraintes connexes.

#### Travail à réaliser :

Dans ce travail pratique, vous devez analyser la structure statistique de textes bruts et déterminer comment ces caractéristiques influencent l'efficacité d'une méthode de compression que vous aurez choisie.

Votre travail sera divisé en deux parties complémentaires :

**Partie 1**— Analyse exploratoire : caractériser les fichiers texte en termes de vocabulaire, fréquences de caractères/mots, entropie, redondance, structure grammaticale ou syntaxique. Ces analyses serviront à formuler des hypothèses sur le comportement attendu des méthodes de compression.

**Partie 2**— Application : choisir et implémenter une méthode de compression sans perte adaptée aux fichiers analysés, en justifiant ce choix à partir des résultats de la Partie 1. Vous évalueriez ensuite ses performances selon différents critères (taux de compression, efficacité, cohérence avec les hypothèses formulées).

Vos réponses aux différentes questions doivent être réunies dans un rapport structuré en format PDF, accompagné du code source complet. Le rapport doit démontrer votre capacité à relier les concepts théoriques du cours aux résultats pratiques obtenus.

## Description des données

Les données sont disponibles sur Moodle et sont réparties comme suit :

- `text1_natural.txt` — extrait de prose littéraire.
- `text2_structured.txt` — données textuelles structurées (CSV).
- `text3_code.txt` — extrait de code source.

Choisissez un fichier de chaque type de données.

## Partie 1 — Analyse exploratoire

### Question 1 (/4)

Décrivez les caractéristiques principales des fichiers texte. Votre analyse doit faire ressortir les éléments qui influencent potentiellement leur compressibilité. Appuyez-vous, lorsque cela est pertinent, sur des figures, visualisations ou mesures quantitatives.

- **2.3 Conceptualiser ou modéliser le problème**  
*Critère d'évaluation : Capacité à extraire des propriétés pertinentes des données textuelles et à formuler clairement les enjeux de leur compression.*

## Partie 2 — Application

### Question 2 (/4)

Quelle méthode de compression avez-vous choisie pour chaque fichier texte ? Justifiez vos choix à l'aide des résultats de l'analyse exploratoire. Formulez des hypothèses testables concernant l'efficacité des méthodes choisies.

- **2.2 Explorer des approches de résolution et planifier la démarche**

*Critère d'évaluation : Justification des choix de votre algorithme au regard des limites rencontrées par celui décrit à la question 1.*

**Question 3** (/8, 4 points par qualité)

Décrivez la démarche expérimentale et les outils utilisés. Présentez vos résultats (taux de compression, rapidité d'exécution, etc.) sous forme de tableaux ou de figures. Expliquez vos choix d'implémentation et les adaptations éventuelles.

- **5.3 Créer ou adapter un outil**

*Critère d'évaluation : Capacité à créer un code informatique fonctionnel.*

**5.4 Intégrer des outils**

Critère d'évaluation : Capacité à bien intégrer les fonctions de bibliothèques de traitement d'images dans son propre code.

**Question 4** (/4)

Discutez de la cohérence entre les résultats observés et les hypothèses formulées. Quelle méthode est la plus adaptée pour chaque type de texte, et pourquoi ?

- **2.5 Valider ses résultats et recommander**

*Critère d'évaluation : Qualité et exhaustivité de l'analyse des résultats.*