

Projet traitement de données

Objectif du projet : Conduire un projet de traitement de données qui permet de mettre en application les connaissances des élèves ingénieur dans le domaine de l'analyse des données et du machine learning à travers une application pratique. Les élèves suivent l'ensemble des étapes d'un processus de projet de machine learning.

Contexte du projet : L'éthique et la moralité en islam touche toute la sphère de vie du musulman et de la communauté musulmane. Ces éthiques musulmanes sont universelles et ce sont des valeurs recherchées par tout être humain dans ses relations avec autrui. Ce projet vient dans un contexte de diffusion des éthiques musulmanes sur les web et les réseaux sociaux.

Problématique : Les éthiques sont qualifiées de bonnes ou bien de mauvaises. Dans le tableau ci-dessous nous donnons une liste des éthiques qui est extraites à partir de plusieurs sources dont des sites web, des livres et des cours qui traitent ce sujet. Nous rencontrons le problème suivant. Nous avons des textes courts en langue arabe. Ces textes sont des citations, des poèmes, des extraits de livres, des sagesses et de l'hadith. Nous voulons associer chacun de ces textes courts à une ou plusieurs éthiques en fonction de son signifié. Nous sollicitons les élèves ingénieurs de nous proposer une solution intelligente et automatique à notre problème en faisant recours aux techniques d'apprentissage automatique.

| الأخلاق الحميدة | | | | |
|---------------------|-------------------|--------------------------|--------------------------------|--------------------|
| الإعراض عن الجاهلين | الاعتدال والوسطية | الإصلاح | الإحسان إلى الغير | الاحترام والتوقير |
| البر | الإيثار | الإنصاف | الأمانة | الألفة |
| التعاون | التضحية | التنبيه | التأني (الأنانة) | البشاشة |
| التوحد | التواضع | التواصي بالخير | التفاؤل | التعافل |
| حسن الظن | حسن السمعة | الحذر والبقرة والحيلة | الجود، والكرم، والسخاء، والبذل | الجديّة والحزم |
| الحياء | الحلم | الحكمة | حفظ اللسان | حسن العشرة والجوار |
| السكينة | السنن | الزهد فيما في أيدي الناس | الرفق | الرحمة |
| الشكر | الشفقة | الشجاعة | السماحة | سلامة الصدر |

| | | | | |
|--|--------------------------|--|--------------------------------------|-------------------------|
| الصَّمْتُ | الصِّلَةُ والتَّوَّاصُلُ | الصِّدْقُ | الصَّبْرُ | الشُّهُامَةُ |
| العَفْوُ والصَّفْحُ | العِفَّةُ | العَزْمُ والعَزِيمَةُ وعلوُّ الهِمَّةِ | العِزَّةُ | الْعَدْلُ والإنصافُ |
| القناعةُ | الفطنةُ والدِّكَاءُ | الفَصَاحَةُ | الفراسةُ | الغَيْرَةُ |
| المُدَاراةُ | المَحَبَّةُ | كَظْمُ الْعَيْظِ | كِتْمَانُ السِّرِّ | القُوَّةُ |
| النَّشَاطُ | النِّزَاهَةُ | المُؤَاسَاةُ | المُزَاخُ | المُرُوءَةُ |
| الْوَفَاءُ | الْوَرَعُ | النَّظَافَةُ | النَّصِيحَةُ | النُّصْرَةُ |
| | | | | الوقارُ والرِّزَانَةُ |
| الأخلاق المذمومة | | | | |
| الإطراء والمدحُ | الإسرافُ والتَّبذِيرُ | الإساءَةُ | الاختِلَافُ والتَّنَازُعُ | الأثرةُ والأنانيَّةُ |
| البُخْلُ والشُّحُّ | الانْتِقَامُ | إِفْشَاءُ السِّرِّ | الإفراطُ | الافتراءُ والنُّبْهَاتُ |
| التَّسَرُّعُ والتَّهَوُّرُ والعَجَلَةُ | التَّخَاذُلُ | التَّجَسُّسُ | البُغْضُ والكِراهِيةُ | البَطَرُ |
| التَّقْلِيدُ والتَّبعيةُ | التَّقْرِيطُ | التَّعَصُّبُ | التَّعْسِيرُ | التَّعَالِي |
| الجَزَعُ | الجدالُ والمراءُ | الجَبْنُ | الثَّرَثَةُ | التَّنَابُرُ بالألقابِ |
| الخِداغُ | الحُبْتُ | الحَقْدُ | الحَسَدُ | الجَفَاءُ |
| الدُّنْ | الدِّيَاثَةُ | الخيَانَةُ | خُلْفُ الوَعْدِ | الجَذْلَانُ |
| سوءُ الجوارِ | السَّفَهُُ والحُمَقُ | السُّحْرِيَّةُ والاسْتِهْزَاءُ | السَّبُّ والشَّتْمُ | رَفْعُ الصَّوْتِ |
| الظُّلْمُ | الطَّمَعُ | الشَّمَاتَةُ | الشَّرَاهَةُ | سوءُ الظَّنِّ |
| العِشُّ | العَدْرُ ونقضُ العهدِ | العُدْوَانُ | العُجْبُ | الغُبُوسُ |
| الفُحْشُ والبِدَاعَةُ | الفُجُورُ | الغِيْبَةُ | الغِلَظَةُ والقِسْوَةُ والْقِظَاطَةُ | الْعَضَبُ |
| اللُّؤْمُ والخِسَّةُ والدَّنَاءَةُ | اللَّامْبَالَاةُ | الكَسَلُ والفُتُورُ | الكَذِبُ | الكِبْرُ |
| تُكْرَانُ الْجَمِيلِ | التَّفَاقُ | الْمَنُ | المَكْرُ والكَيْدُ | المُدَاهَنَةُ |
| اليأسُ والفُتُوْطُ والإحباطُ | الْوَهْنُ | الهُمَزُ واللُّمَزُ | الهُجْرُ | النَّمِيْمَةُ |

Conduite du projet traitement de données :

Nous allons suivre un ensemble des étapes d'un processus de machine learning :

1. Définition du problème et le contexte du projet
2. Sélection et collecte des données
3. Description des données
4. Exploration des données
5. Préparation des données
6. Représentation et transformation des données
7. Sélection des "features"
8. Sélection de modèles d'apprentissage
9. Apprentissage
10. Optimisation
11. Evaluation du modèle
12. Déploiement du modèle

Tâche 1 : Définition du problème et le contexte du projet : décrire le contexte de votre projet et le problème.

Tâche 2 : collecte de données et constitution de corpus

1. Tous les textes collectés doivent être en arabe littéraire.
2. Dresser une liste des bonnes éthiques et des mauvaises éthiques (الأخلاق الحميدة و الأخلاق المذمومة)
3. Caractériser chaque éthique que ce soit bonne ou mauvaise par une ou plusieurs définitions ainsi que des synonymes.
4. Collecter des textes courts (sagesse, proverbe, citation) ayant pour sujet une éthique particulière. Faire ceci pour toutes les éthiques que ce soit bonnes ou mauvaises.
5. Un seul texte court peut parler de plusieurs éthiques.

Bon courage