Test des algorithmes d’apprentissage machine pour CRCLASS:

- Base de données:
  - Swift: 2000 individus concernant 2 classes(GRB et Bruit).
  - Fermi: 6400 individus concernant 6 classes(GRB, LOCLPAR, SFLARE, SGR, TGF, TRANSNT)
- Méthodes:
  - Retenue : Random Forest
  - Testées : AFD, ACP, ADL, K-means, SVM
- L’entraînement d’un modèle de classification pour la base de données sur Fermi:
  - Entrée du chaque individu:
    - 9 paramètres: lii, bii, trigger_timescale, trigger_algorithm, channel_low, channel_high, detector_mask, geo_long, geo_lat
    - 1 étiquettes : GRB, LOCLPAR, SFLARE, SGR, TGF ou TRANSNT
  - Sortie: Pourcentage de la probabilité pour chaque individu à prédire

    Exemple:
    GRB : 90.3%
    LOCLPAR : 5.0%
    SFLARE : 0.2%
    SGR : 0.0%
    TGF : 0.0%
    TRANSNT : 4.5%

Résultat:

- Performance du modèle de Random Forest pour la base de données Swift: 90%, Fermi: 89%

- Exemples:

bn121010862(TGF):
GRB :0.0%
LOCLPAR :0.2%
SFLARE :0.0%
SGR :0.0%
TGF :99.8%
TRANSNT :0.0%

bn121231445(GRB):
GRB :80.8%
LOCLPAR :1.9%
SFLARE :17.1%
SGR :0.0%
TGF :0.0%
TRANSNT :0.2%

bn150416992(LOCLPAR):
GRB :30.2%
LOCLPAR :60.1%
SFLARE :5.9%
SGR :0.0%
TGF :0.0%
TRANSNT :3.8%

bn150301312(SGR):
GRB :40.4%
LOCLPAR :0.3%
SFLARE :0.3%
SGR :52.7%
TGF :5.5%
TRANSNT :0.8%

- Exemples mal classés:

bn120204054(GRB):
GRB :45.2%
LOCLPAR :3.3%
SFLARE :50.3%
SGR :0.0%
TGF :0.0%
TRANSNT :1.2%

bn081120618(GRB):
GRB :18.0%
LOCLPAR :1.0%
SFLARE :80.8%
SGR :0.0%
TGF :0.0%
TRANSNT :0.2%

bn140201766(SFLARE):
GRB :79.6%
LOCLPAR :0.0%
SFLARE :20.2%
SGR :0.0%
TGF :0.0%
TRANSNT :0.2%

bn141025734(SFLARE)
GRB :13.4%
LOCLPAR :61.9%
SFLARE :24.2%
SGR :0.1%
TGF :0.0%
TRANSNT :0.4%

Types d’événement en sortie de CRCLASS:

- Catégories:
  - SGRB: Short Gamma-ray Burst
  - LGRB: Long Gamma-ray Burst
  - LOCLPAR: Local Particles
  - SFLARE: Solar Flare
  - SGR: Soft Gamma Repeater
  - TGF: Terrestrial Gamma-Ray Flash
  - TRANSNT: Generic Transient

Données simulées pour base d’apprentissage SVOM:

- GRB
- Autres classes
- Il faudrait idéalement un minimum de 1000 exemples dans chaque classe pour l’entraînement.