

Python data science

Nicolas Rousset

Formation Data Science, day 1

Python lui-même est lent

- Avoir un interpréteur / traducteur ralentit le processus
- Le typage dynamique force le programme a passer beaucoup de temps a comprendre ce que vous voulez faire

$c = a + b$

- Le typage dynamique ne permet pas d'avoir des structures de données efficaces

Structure de données efficaces



NumPy array
100x / 1000x faster



Python array
Slow

- pour l'analyse d'image ou de donnée, pour la performance il n'y a pas de salut en dehors des langages compilés (c'est à dire du C ou C++ 99% du temps actuellement)
- les librairies / programmes issus de code compilé dépendent des systèmes d'exploitation
- python s'interface très bien avec les langages compilés
- on utilise python comme un manager de haut niveau => il ne fait pas de calcul, il dit aux librairies C++ quoi faire

Il n'y a pas de magie en informatique

- Utiliser python vous permet de faire des études avec du code haut niveau et une efficacité de code bas niveau ...
- Mais de temps en temps il est nécessaire de comprendre la structure bas-niveau sous-jacente
- Python a perdu beaucoup de compatibilité windows / linux dans le process
- Vous pouvez utiliser sans problème python sur l'un ou l'autre des systèmes mais les deux pour un même projet peut être risqué