

11 décembre 2023

Conférence annuelle XBRL France
«REPORTINGS : CAP 2024 »

avec le soutien de la  CNCC



iXBRL et machine learning

- Marc Houllier
- Corporatings
- XBRL France

Tâche



Texte en italien



Texte en anglais

Flux vidéo



Rotation du volant

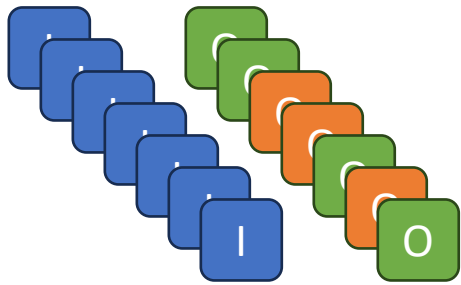
Fragment de rapport



Signification du fragment



Apprentissage supervisé ou non-supervisé



Score **non-supervisé**

Reproduire l'entrée

Score : distance entre I and O



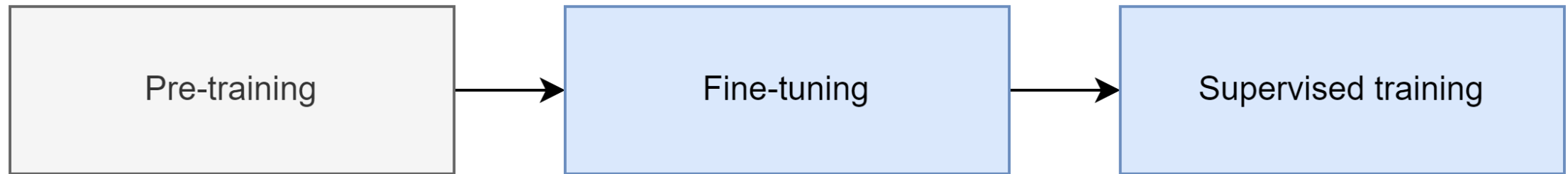
Score **supervised**

Dire si l'image contient un chien

Score : Correspondance avec étiquetage



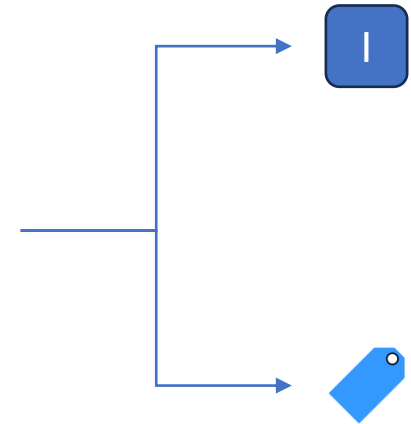
Apprentissage supervisé ou non-supervisé



Apprentissage supervisé sur texte

Trouver les trois mots les plus importants dans un texte, dans l'ordre.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod **<imp order="2">tempor</imp>** incididunt ut labore et dolore magna aliqua. Ut enim ad **<imp order="1">minim</imp>** veniam, quis nostrud exercitation ullamco laboris nisi ut **<imp order="3">aliquip</imp>** ex ea commodo consequat.



Apprentissage supervisé sur rapport financier

In 2022, an expense of (<ix:nF name="ExpensesFromShareBasedPayments" ...>1</ix:nF>) million euros (including social security contributions) was recognized with corresponding entries in equity (<ix:nF ...>10</ix:nF> million euros) and employee benefits (<ix:nF ...>1</ix:nF> million euros).

Tâche : Apprendre à identifier une caractéristique d'un passage de texte.

Applications : classifications

- Classification de valeur
 - En isolation, ou à partir d'un document précédent
 - Classification d'extension
 - Classification de table
 - Classification de bloc de texte
-

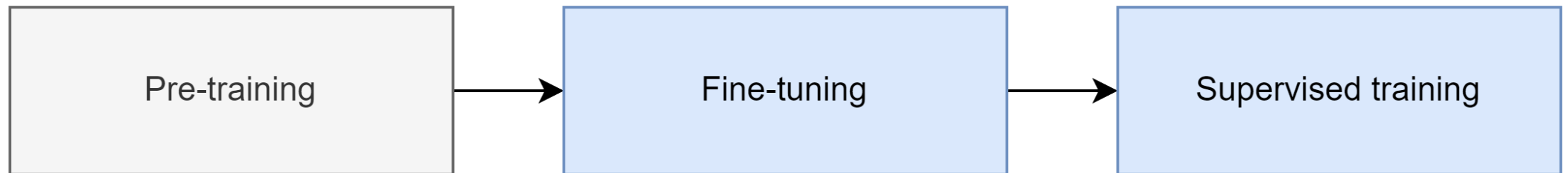
Applications : détection d'erreur

- Détection d'anomalies
- Détection d'éléments utilisés de façon équivalente

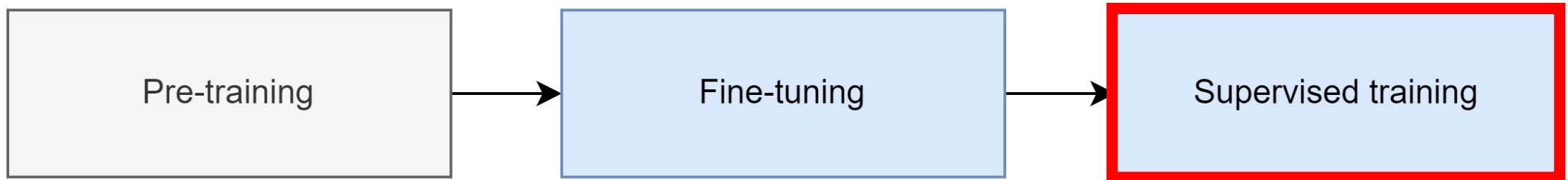


Quantité de données d'entraînement

- Français: 400 rapports par an
- Lithuanien: 11 rapports par an



Quantité de données d'entraînement



Quel que soit le progrès des modèles généralistes (pré-entraînement), les ensembles d'entraînements étiquetés seront toujours nécessaires !

