



UNIVERSITATEA TEHNICĂ

DIN CLUJ-NAPOCA

FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE
SPECIALIZAREA CALCULATOARE

Proiect TensorFlow

Urcan Denisa-Teodora
Calculatoare, grupa 30241

1. Definitie

Un algoritm machine learning (ML) este un program ce folosește inițial cantități mari de date, ca mai apoi să poată face "predicții" despre datele ce vor apărea ulterior în sistem. TensorFlow este o bibliotecă folosită în cadrul învățării progresive (deep learning), dezvoltată de Google și care excellează în calculele matematice. Detine, de asemenea, un set bogat de interfețe de programare (API) în majoritatea mediilor limbajelor de programare pentru proiectele de învățare progresivă. Astfel de exemple de medii de programare sunt Python, C, C++, Rust, Haskell, Go, Java, Android, iOS, Mac OS, Windows, Linux și Raspberry Pi.

Unitatea centrală folosită de această bibliotecă este un tensor, care se definește printr-un set de valori primitive modelate sub forma unui vector de dimensiune variabilă. Acest număr mare de vectori reprezintă motivul pentru care procesoarele a căror responsabilitate sunt calculele matematice a numerelor reprezentate în virgula mobilă excellează în ceea ce privește scăderea vitezei de execuție a acestor algoritmi. Un tensor este un set de primitive dintr-o matrice multidimensională, clasificate după numărul de dimensiuni. Exemple de tensor pot fi:

0: 5.0

1: [5.0, 10.0, 15.0, 20.0, 25.0]

Ca multe biblioteci celebre, TensorFlow este scrisă în C++, iar cel mai popular port este pentru Python. Această bibliotecă respectă mai fidel filozofia limbajului de programare Python.

Pentru a memora tensorii cu scopul utilizării acestora, se folosesc variabile, care pot fi și statice (constante). O remarcă importantă care trebuie menționată este necesitatea inițializării acestor variabile înainte de folosirea lor. Bibliotecă construiește un arbore al operațiilor care trebuie realizate, iar în momentul apelării unei "sesiuni" arborele începe să ruleze. Se folosește acest tip de structură datorită scalabilității, care se reflectă prin schimbarea reprezentării interne a datelor în tensori, numiți și vectori multi-dimensionali.

Un program TensorFlow este structurat într-o fază de construcție, care assemblează arborele de date și o fază de execuție, care execută operațiile în arbore.

Această bibliotecă este folosită în cadrul multor companii și este o componentă open-source. Arhitectura sa paralelă și abilitatea accesării datelor din Hadoop HDFS, sistem de fișiere distribuit care se ocupă de managementul seturilor mari de date.

TensorFlow este cunoscută drept o bibliotecă pentru calcule numerice ce folosește grafuri cu fluxuri de date. Datele de intrare sunt transformate în flux de date prin aceste grafuri non-ciclice (arbori), care presupune o schimbare de paradigmă în programare, orice date putând fi reprezentate printr-un graf în care nodurile sunt calcule, iar liniile ce unesc nodurile sunt numere. Calculele sunt numite operatori, iar fluxul de date pe linii este numit tensors.

TensorFlow excelează în mediile de producție datorită arhitecturii distribuite și paralel și a capacității sale de a accesa date. O altă caracteristică cheie a TensorFlow este TensorFlow Serving, care este un server de producție de înaltă performanță pentru implementarea și deservirea modelelor de învățare automată folosind gRPC ca protocol de rețea.

Rețelele neuronale există pe piața de multă vreme, dar cu ajutorul acestei librării pot fi rulate într-un timp rezonabil datorită computerelor mai puternice și a GPU-urilor.

Un exemplu de implementare a unei rețele neuronale ar fi definirea unei arhitecturi, încărcarea datelor în model, împărțirea datelor în loturi destinate preprocesării, convertirea acestora și utilizarea acestora pentru învățare. Pe măsură ce modelul este antrenat, modelul este salvat pentru reutilizare. Execuțiile viitoare testează modelul pe date noi pentru a verifica performanța. Acest lucru este ajutat de rețele rapide, procesoare rapide, GPU-uri rapide, RAM mare, hard disk-uri mari și, în general de computerele mai rapide.

2. Arhitectura

La nivel de arhitectura, TensorFlow este o librerie structurată pe 3 nivele, printre care se numără librăriile de nivel înalt, nucleu și platforma.

3. Cazuri de utilizare

În ceea ce privește utilizarea sa, se remarcă în:

- recunoașterea vorbirii, prin identificarea a ce spun oamenii și convertirea în text,
- recunoaștere de imagine,
- analiza sentimentelor,
- sumarizarea textului,
- procesarea imaginilor și a videoclipurilor provenite de la un dispozitiv mobil,
- detectia obiectelor în imagini,
- adăugarea subtitrării și a posibilelor paragrafe care descriu o imagine,
- chatbot ca și răspuns inteligent continuând replici generate,
- detectia cancerului prin compararea imaginilor faciale sau corporale.

Una dintre cele mai relevante aplicații ale acestei librării se referă la recunoașterea de imagini. Se remarcă prin portabilitate, putând fi folosită pe sisteme de operare diferite, printre care se menționează Mac, Raspberry Pi, telefon mobil.

O altă aplicație este în cadrul rețelelor neuronale, care cu ajutorul acestei librării și datorită dezvoltării calculatoarelor pot rula într-un interval de timp rezonabil.

4. Exempleu de cod

```
import tensorflow as tf

firstnumber = tf.constant([10, 20, 30, 35, 40, 50])
secondnumber = tf.Variable(firstnumber + 50)

with tf.Session() as session:
    session.run(tf.global_variables_initializer())
    print(session.run(secondnumber))
```