



Python, C++ или Matlab за изкуствен интелект

В темата ще бъдат разгледани следните основни въпроси:

- Развитие на трите езика: Python, C++ и Matlab
- Машинно обучение с Python, C++ и Matlab
- Дълбоко обучение с Python, C++ и Matlab
- Сравнение между езиците за програмиране Python, С ++ и Matlab по 13 основни характеристики

------ <u>www.eufunds.bg</u> ------





1. Развитие на трите езика

1.1. Кратка история на Matlab

Маtlab, което е съкращение от "Мatrix Laboratory", е патентована платформа за програмиране, предназначена за създаване и анализиране на системи и продукти от инженери и учени. Първоначално тя не е създадена като език за програмиране, а като прост интерактивен матричен калкулатор. Matlab е търговски продукт и се предлага със собствена работна среда. Исторически това се е разглеждало като предимство поради факта за поддръжка на системата при възникнали проблеми. В наши дни почти всички езици за програмиране, използвани за разработки, са с отворен код и са безплатни.

Маtlab не е продукт с отворен код, което означава, че разработчиците нямат достъп до вътрешната работа на неговия патентован код. Той доминира в академичните среди, където изчислителната му мощност го прави популярен за цифрова обработка на сигнали.

Първоначално разработен от математика и компютърен програмист Клив Молер през 60-те години на миналия век, той не го продава, а го разпространява безплатно в университетите. Едва през 80-те години на миналия век е пуснат като комерсиален продукт. Затова може да бъде изключително скъпо за обикновените потребители да закупят дори един негов лиценз.

------www.eufunds.bg ------





1.2. Кратка история на Python

Руthon е интерпретиран, обектно-ориентиран език за програмиране на високо ниво с динамична семантика и с общо предназначение. Неговите вградени структури от данни на високо ниво, комбинирани с динамично въвеждане и динамично обвързване, го правят много привлекателен за бързо разработване на приложения, както и за използване като скриптов или език за свързване на съществуващи компоненти. Простият, лесен за научаване синтаксис на Руthon подчертава четливостта и следователно намалява разходите за поддръжка на програмата. Руthon поддържа модули и пакети, което насърчава програмната модулност и повторното използване на кода. Интерпретаторът на Руthon и общирната стандартна библиотека са налични в изходен код или двоична форма безплатно за всички основни платформи и могат да се разпространяват свободно. Той е с отворен код, което означава, че всеки може да получи достъп до него, да се рови в кода му, за да види как работи и дори да създаде свои собствени пакети за него. Това е едно от основните предимства на Руthon пред Matlab.

Руthon е създаден в края на 80-те години и е внедрен за първи път през декември 1989 г. Оттогава съществува, става все по-популярен и постоянно се нарежда като един от най-популярните езици през годините. Има огромна и нарастваща функционалност. Добър избор е както за прости задачи, от които

------ <u>www.eufunds.bq</u> ------





може да се интересуват начинаещите, така и за сложни задачи, от които се интересуват корпорациите и организациите.

1.3. Кратка история на С++

Разработването на програмния език C++ от датския програмист Бярне Строуструп в лабораториите на Бел започва през 1979 г. Тъй като C++ е опит за добавяне на обектно-ориентирани функции (плюс други подобрения) към С, през първите години той се е наричал "С с обекти". С развитието на езика Страуструп го преименува като C++ през 1983 г. За първи път C++ се стандартизира през 1998 като ISO/IEC 14882:1998, през 2003 година стандарта се преразглежда като ISO/IEC 14882:2003, а от 2011 стандартът се обновява на всеки три години.

Независимо от своя по-сложен синтаксис в сравнение с другите програмни езици С++ е добър избор за изкуствен интелект, защото той предлага най-голямо бързодействие. Библиотеките му за машинно и дълбоко обучение са ограничени на брой, но са изключително ефективни. Езикът С++ предоставя библиотеките SHARK и MLPACK за интегрирани ML решения.

2. Машинно обучение с Python, C++ и Matlab

Руthon разполага с множество библиотеки и пакети както за старите, така и за новите модели на машинно обучение. Освен това Python е найшироко използваният език за съвременни изследвания на машинно обучение

----- <u>www.eufunds.bg</u> ------





в индустрията и академичните среди. Мани Тадайон казва в статията си: "Машинното обучение е областта, в която Python и R имат ясно предимство пред Matlab."

Маtlab може да се използва за машинно обучение. Мatlab разполага с голяма колекция от патентовани алгоритми, които могат да предложат незабавна визуална обратна връзка. Основният недостатък на Matlab е малката преносимост на кода.

Matlab може да се използва за машинно обучение, но Python е едновременно по-добър и по-преносим в тази област.

С++ библиотеката SHARK съдържа бързи модули, поддържащи алгоритми за обучение с надзор, като линейна регресия, невронни мрежи, групиране, k- стойности и т.н. Тя също така включва функционалността на линейната алгебра и числената оптимизация, които са много важни при изпълнение на задачи за машинно обучение.

МLРАСК библиотеката е бърза и гъвкава библиотека за машинно обучение, написана на С++. Тя цели да осигури бързи и разширяими реализации на авангардни алгоритми за машинно обучение, като предоставя тези алгоритми като прости програми от командния ред. Библиотеката предоставя възможност за свързване на код от Python и Julia със С++ класове, които след това могат да бъдат интегрирани в по-мащабни решения за машинно обучение.

------ www.eufunds.bg ------





Shogun е библиотека за машинно обучение с отворен код, която предлага широка гама от ефективни и унифицирани методи за машинно обучение. Библиотеката е реализирана на С++ и предоставя автоматично генерирани, унифицирани интерфейси към Python, Octave, Java/Scala, Ruby, С#, R, Lua. Shogun предоставя лесна комбинация от множество представяния на данни, класове, алгоритми и инструменти с общо предназначение за бързо създаване на прототипи на канали за данни.

DyNet (Dynamic Neural Network Toolkit) е библиотека за невронни мрежи, написана на С++ и обвързвания в Python. Тя е проектирана да бъде ефективна, когато работи на СРU или GPU. DyNet изгражда своя изчислителен граф в движение, което прави моделите с променлив вход и променлив изход лесни за изпълнение с висока производителност. Библиотеката е много подходяща за техники като обучение с утвърждаване, обработка на естествен език, графични структури и други.

Написана на C++, OpenNN (Open Neural Networks) е библиотека за невронни мрежи с отворен код за разширен анализ. Библиотеката съдържа усъвършенствани алгоритми и помощни програми за работа със следните решения за изкуствен интелект, като класификация, регресия, прогнозиране и други. Основното предимство на тази библиотека е нейната висока производителност.

Бързата изкуствена невронна мрежа FANN (Fast Artificial Neural Network) е библиотека за невронни мрежи с отворен код, написана на език С.

------ <u>www.eufunds.bq</u> ------





Библиотеката прилага многослойни изкуствени невронни мрежи на С като поддържа както за напълно свързани, така и слабо свързани мрежи. Тя е лесна за използване, многофункционална, добре документирана и бърза. Функциите включват обучение с обратно разпространение, обучение на развиваща се топология, крос-платформа и др. Библиотеката може да използва както числа с плаваща, така и такива с фиксирана запетая.

3. Дълбоко обучение с Python, C++ и Matlab

Дълбокото обучение се счита за еволюция на машинното обучение, което използва програмируема невронна мрежа, позволяваща на машините да вземат точни решения без помощта на хората. Като такова дълбокото обучение може да се счита за подмножество на машинното обучение.

Теоретично, моделите за дълбоко обучение трябва да могат сами да определят дали прогнозите са точни, докато машинното обучение изисква повече човешко ръководство. Това е част от причината да се смята, че Python за най-добрия вариант за дълбоко обучение.

Това, че Matlab е патентован, означава, че иновациите се въвеждат побавно и с допълнително заплащане. Python е с отворен код и всеки може да създава пакети, които стават широко възприети от общността на Python. Пакети като PyTorch, Tensorflow, Caffe и т.н. се използват широко за дълбоко обучение. Повечето онлайн курсове за дълбоко обучение включват Python като език за програмиране.

------ www.eufunds.bg ------





В обобщение, Python е по-добър от Matlab за дълбоко обучение. Matlab е по-бърз от Python, но Python е по-добър в паралелното изпълнение на множество задачи.

Интересно е да се отбележи, че повечето посочени библиотеки за дълбоко обучение не са написани само на Python, а голяма част от тях са написани на C++.

ТепsorFlow е библиотека за внедряване на модели на мобилни, микроконтролери и други крайни устройства. В по-голямата си част ядрото на TensorFlow не е написано на Python, а в комбинация от силно оптимизиран С++ и CUDA, кйто е език на Nvidia за програмиране на графичен процесор GPU (Graphics Processing Unit).

Конволюционна архитектура за бързо вграждане на характеристики Caffe (Convolutional Architecture for Fast Feature Embedding) е рамка за дълбоко обучение, написана на C++. Характеристиките на тази библиотека включват експресивна архитектура, разширим код, скорост и голяма общност, която насърчава активното развитие в научните изследвания и внедряването в индустрията.

Когнитивният инструментариум за дълбоко обучение на Майкрософт CNTK (CogNitive ToolKit) е написан на C++, който описва невронните мрежи като поредица от изчислителни стъпки чрез насочен граф. Той прилага обучение със стохастичен градиент SGD (Stochastic Gradient Descent) и обратно разпространение на грешки, използвайки автоматично

------ <u>www.eufunds.bq</u> ------





диференциране и паралелизиране в множество GPU и сървъри. CNTK позволява на потребителите лесно да реализират и комбинират популярни типове модели като дълбоки невронни мрежи за обучение DNN (Deep Neural Network) с подаване напред, конволюционни мрежи (CNN) и рекурентни мрежи RNN RNN (Recurrent Neural Network) или LSTM (Long Short-Term Memory).

4. Сравнение между езиците за програмиране Python, С ++ и Matlab

Ключовите разлики между Python, С ++ и Matlab, които са важни от гледна точка на езиците за програмиране, са обсъдени по-долу и накратко са обобщени в таблица 2.1. Сравнението е направено по 13 основни характеристики: компилация, употреба от начинаещи, естество на езика, преносимост, събиране на боклук, типове, обхват на променливите, бързо създаване на прототипи, функции, ефективност и сложност на синтаксиса, скорост на изпълнение, производителност и популярност

Компилация

С ++ е език използващ компилатор. Компилаторът на С++ генерира обектен код от изходния код на С++ и след това той се изпълнява, за да произведе изходния изпълним код.

------ <u>www.eufunds.bg</u> ------





Python е език използващ интерпретатор. Кодът на Python с разширение ру не е необходимо да се компилира. Може директно да се предаде на интерпретатора на Python и да се генерира изхода.

Аналогично на Python, Matlab е език използващ интерпретатор.

Употреба от начинаещи

С ++ има много функции и също така има сравнително труден синтаксис, което го определя като не толкова лесен език за начинаещи.

Python е лесен за писане и има ясен синтаксис. Следователно писането на програми на Руthon е много по-лесно в сравнение с С ++.

Писането на код на Matlab отново не е много лесна задача, защото и той поддържа много функции и има сравнително труден синтаксис.

Естество на езика

С ++ е статичен език, т.е. декларацията на променливите, типът данни на променливите и т.н. се проверяват по време на компилиране. Това поддържа изходния код без грешки по време на изпълнение.

Руthon, от друга страна, не е статично типизиран. По време на компилиране не се извършва проверка на типа. Следователно кодът е податлив на грешки. Аналогично на Python и Matlab използва интерпретатор, което отново води до поява на грешки по време на изпълнение на програмите.

------ <u>www.eufunds.bg</u> ------





Преносимост

С ++ не е преносим, т.е. написаният веднъж код трябва да се прекомпилира на всяка различна платформа.

Python е преносим. Освен това е междуплатформен и програмите могат да се изпълняват на всяка платформа.

Matlab, аналогично на C ++, е непреносим.

Събиране на отпадъци/управление на паметта

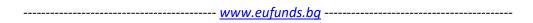
В С ++ управлението на паметта е ръчно. С ++ не поддържа автоматично събиране на отпадъци от ресурси.

Python, от друга страна, има функция за автоматично събиране на отпадъци. Управлението на паметта му се контролира от системата.

В Matlab управлението на паметта отново е ръчно с изключение на освобождаването на хип паметта при излизане от функция. МАТLAB се опитва да освободи хип паметта всеки път, когато стекът се освобождава, т.е. при излизане от всяка функция. Това има доста предимства, но води до забавяне на изпълнението, особено ако освобождаването на хипа е доста бавно.

Бързо създаване на прототипи

С++ не позволява бързо прототипиране на код.







Използвайки Python, може да се направи бързо прототипиране на код, така че да може да се използва по-късно за изграждане на приложения, използващи езици от по-високо ниво. В Matlab е възможно прототипиране само за някои FPGA платформи.

Обхват на променливите

С ++ има код, маркиран от блокове с помощта на фигурни скоби {} и цикли. Обхватът на променливите е ограничен до тези блокове и цикли, обозначени с {}.

Обхватът на променливите, използвани в Python, не е ограничен до блокове или цикли. Променливите са достъпни дори извън фигурните скоби.

В Matlab променливите, декларирани в дадена функция, имат ограничена видимост от други функции. Обикновените променливи са валидни в работното пространство.

Типове

В С++ типовете данни са обвързани с имена и се проверяват по време на компилиране. Това намалява възможността за всякакви грешки по време на изпълнение.

В Python типовете данни са обвързани със стойности и се проверяват по време на изпълнение. Кодът може да е по-податлив на грешки по време на изпълнение, тъй като тези грешки не се улавят по време на компилиране.

------ www.eufunds.bg





В Matlab типовете данни са обвързани с имана и стойности. Те се проверяват по време на изпълнение и затова кодът може да е по-податлив на грешки по време на изпълнение.

Функции

Функциите са блокове от кодове с един или повече параметри и върната стойност. Всеки от параметрите и връщаната стойност има тип.

В С++ типовете параметри и типът на връщане по време на извикване на функция трябва да съвпадат с тези в дефиницията на функцията.

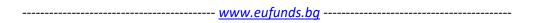
В Python няма такова ограничение за типовете параметри и връщани стойности.

В Matlab отново няма ограничение за типовете параметри и връщани стойности на функциите. Ако е необходимо да се въведат ограничения в използваните параметри, то те се дефинират със специална функция за валидация.

Ефективност и сложност на синтаксиса

С ++ кодът е труден за поддръжка, тъй като може да стане сложен за четене, когато решенията стават по-големи.

Руthon, от друга страна, има чист код и прост синтаксис. Изходният код за Python е по-лесен за поддръжка.







Кодът на Matlab е сложен за четене при големи проекти и като такъв е труден за поддръжка.

В С ++ има ясно разграничаване на кода чрез използване на блокове, затворени в {}, точка и запетая, показваща края на израза и т.н. Така в С++ синтаксисът е добре организиран.

В Python няма блокове или точка и запетая. Вместо това Python използва отстъп.

В Matlab има ясно разграничаване на кода чрез използване на блокове, ограничени с даден оператор и завършващи с ключовата дума end.

Скорост на изпълнение

Що се отнася до скоростта на изпълнение, програмите на С ++ работят най-бързо. Всъщност С ++ е известен и се използва широко в приложения, които трябва да работят бързо, особено в платформи за игри.

Python, от друга страна, работи бавно, по-бавно от програмите на Java. Следователно Python може да се използва в приложения, които нямат специални изисквания към скоростта.

Производителност

С ++ е статичен език, което елиминира грешките по време на изпълнение. С ++ също създава по-стабилен и по-бърз код за изпълнение. Това прави С++ език с висока производителност.

------ www.eufunds.bg ------





Тъй като Python е динамичен език, то съществува възможност някои грешки или нежелана ситуация да възникнат по време на изпълнение на програмите.

Аналогично на Python, Matlab е динамичен език, и следователно съществува възможност някои грешки или нежелани ситуации да възникнат по време на изпълнение на програмите.

Така че, що се отнася до производителността, Python и Matlab изостават от С ++. Но когато става въпрос за машинно и дълбоко обучение, Python е този, който има определени предимства спрямо С++ и Matlab.

Популярност

Руthon е лесен за научаване и лесен за прилагане на практика в сравнение със С ++ и Matlab. Друго предимство на Руthon са неговите библиотеки, които позволяват да се пише кратък и лесно четим код прилагащ готови функции, особено тези за анализ на данни и машинно обучение.

Това е особено полезно, когато се разработват сложни приложения за изкуствен интелект прилагащи машинно и дълбоко обучение.

По отношение на популярността Python е избор номер едно за програмистите, особено за разработването на приложения за машинно и дълбоко обучение, това. Езикът С++ със своето най-голямо бързодействие е най-популярен за вградени или корпоративни приложения. Matlab остава най-добрия избор при обработка на изображения и обработка на сигнали.

------ www.eufunds.bg





Таблица 2.1. Сравнение между езиците за програмиране Рython, С ++ и Matlab

Характеристика	Python	C ++	Matlab
Компилация	Интерпретатор	Компилатор	Интерпретатор
Употреба от	По-лесно писане на	Не е лесно да се	Не е лесно да се
начинаещи	код	пише код	пише код
Естество на езика	Динамичен	Статичен	Динамичен
Преносимост	Преносим	Не е преносим	Не е преносим
Събиране на	Поддържа	Не поддържа	Не поддържа
боклук			
Типове	Типове данни,	Обвързани със	Обвързани със
	обвързани с имена,	стойности,	стойности,
	проверени по време	проверени по	проверени по
	на компилиране	време на	време на
		изпълнение	изпълнение
Обхват на	Достъпни извън	Ограничен в	Достъпни извън
променливите	циклите или	рамките на	циклите или
	блоковете	циклите или	блоковете
		блоковете,	
		динамични	
		проенливи	
		достъпни извън	
		цикли и блокове	
Бързо създаване	Възможно	Не е възможно	Не е възможно
на прототипи			
Функции	Няма ограничения	Ограничения за	Ограниченията в
	за вида на	типа на	използваните
	параметрите или	параметрите	параметри се
	връщаната стойност	или връщаната	дефинират със
		стойност	специална
			функция за

-----<u>www.eufunds.bg</u> ------





			валидация
Ефективност и	Труден за	Лесен за	Труден за
сложност на	поддръжка.	поддръжка.	поддръжка.
синтаксиса	Без блокове или	Използва	Използва
	точка и запетая	блокове и точка	блокове
		и запетая	
Скорост на	бавен	Най-бърз	бавен
изпълнение			
Производителност	Ниска	Висока	Ниска
Популярност	Най-популярен за	Най-популярен	Най-популярен
	машинно обучение	за вградени или	за обработка на
	и дълбоко обучение	корпоративни	изображения и
		приложения	обработка на
			сигнали

Обобщавайки гореизложеното, всеки програмист може да определи кой език за програмиране е най-подходящ за следващия му проект с изкуствен интелект, като изборът му ще зависи от това, за какво се отнася идеята му за проект, както и от опита му с различни езици за програмиране.

------www.eufunds.bg ------