Is python the best language to learn Machine Learning?

Lets start with knowing the pros and cons of using python for machine learning

Pros:

- Tons of open source code

- Great data analysis libraries (pandas[2], statsmodels[3])

- Great visualization libraries (matplotlib[4],ggplot[5])

- Code can be easily integrated into a web service written in Python (Flask[6], Django[7], Pyramid[8], etc)

Cons:

- A lot of cutting edge, advanced academic research is still being done in R/Matlab

In machine learning you want at least two things:

Effective implementations

Simplified API

Most of the time authors of ML frameworks develop them in C/C++/Java and then provide the interface in Python. Truth is, for most (99%) ML tasks you don't need to reimplement anything yourself, so the existing well-written library which happens to be in Python will do just fine. That propagates the growth of Python-centered stack thus making Python the most convenient language for ML.

It's really amazing how much can be achieved with so little coding.

Python is open source, it has a huge user community, and a number of good packages for machine learning.

Matplotlib/pylab are packages that offer good solid visualization, IMO comparable to R’s visualization. The design of Python stresses good code readability.

It is a full fledged language offering simple objected oriented design. The only downside to Python is its execution speed. In many cases, this can be overcome by careful optimization of the data structures.

Various efforts have been made to improve the speed of Python. Pypy is a just-in-time compiler for Python that can sometimes speed up your Python code, depending on what it does. Cython allows for low level optimization using C by first translating Python code to equivalent C code.

For beginners, yes, Python is the best choice. I wouldn't worry too much about the functional programming aspect though. Yes, occasionally you will be using lambdas, map, reduce (fold/fold sorts, in terms of Haskell) but I wouldn't expect that to become anything more than just a decent side kick. Explore libraries like scikit-learn and Shogun (the latter is lesser known to the Python community, but the coolest thing is that it, using some SWIG magic, provides APIs in pretty much every language you'd ever code for ML).

Go for python if you want to use libraries for various algorithms. If you are going for text parsing, nlp then it is better to use python as it has many free libraries like nltk, textblob etc which can be downloaded and used easily.

If you want to program your own algorithm with only numbers like some regression or classification algorithm then it is best to go with MATLAB or OCTAVE where you could actually control all the optimizations by fixing various regularizations parameters and can add on your own.

Python leads the pack, with 57% of data scientists and machine learning developers using it and 33% prioritizing it for development. Little wonder, given all the evolution in the deep learning Python frameworks over the past 2 years, including the release of TensorFlow and a wide selection of other libraries.

Our data reveals that the most decisive factor when selecting a language for machine learning is the type of project you’ll be working on — your application area. In our survey we asked developers about 17 different application areas while also providing our respondents with the opportunity to tell us that they’re still exploring options, not actively working on any area. Here we present the top and bottom three areas per language: the ones where developers prioritise each language the most and the least.

Machine learning scientists working on sentiment analysis priorities Python (44%) and R (11%) more and JavaScript (2%) and Java (15%) less than developers working on other areas. In contrast, Java is prioritized more by those working on network security / cyber attacks and fraud detection, the two areas where Python is the least prioritized. Network security and fraud detection algorithms are built or consumed mostly in large organizations — and especially in financial institutions — where Java is a favorite of most internal development teams. In areas that are less enterprise-focused, such as natural language processing (NLP) and sentiment analysis, developers opt for Python which offers an easier and faster way to build highly performing algorithms, due to the extensive collection of specialized libraries that come with it.

Artificial Intelligence (AI) in games (29%) and robot locomotion (27%) are the two areas where C/C++ is favored the most, given the level of control, high performance and efficiency required. Here a lower level programming language such as C/C++ that comes with highly sophisticated AI libraries is a natural choice, while R, designed for statistical analysis and visualizations, is deemed mostly irrelevant. AI in games (3%) and robot locomotion(1%) are the two areas where R is prioritized the least, followed by speech recognition where the case is similar.

Other than in sentiment analysis, R is also relatively highly prioritized — as compared to other application areas — in bioengineering and bioinformatics (11%), an area where both Java and JavaScript are not favored. Given the long-standing use of R in biomedical statistics, both inside and outside academia, it’s no surprise that it’s one of the areas where it’s used the most. Finally, our data shows that developers new to data science and machine learning who are still exploring options priorities JavaScript more than others (11%) and Java less than others (13%). These are in many cases developers who are experimenting with machine learning through the use of a 3rd-party machine learning API in a web application.

COCNCULSION:

If we go through the Analytics and the how easy python is as compared to other langue then yes python is the best langue to learn Machine Learning.