anadrol

053.314.328-40 Identificação

custom\_dictionaries\_only

custom\_terms\_only

Jane Smith

5446882375129691

Exp. 09/16

John Smith

4532014495343

Exp. 02/17A question I get asked a lot on my email list is: what is the best programming language for machine learning?

I’ve replied to this question many times now it’s about time to explore this further in a blog post. Ultimately, the programming language you use for machine learning should consider your own requirements and predilections. No one can meaningfully address those concerns for you.

What Languages Are Being Used

Before I give you my opinion, it is good to have  a look around to see what languages and platforms are popular in self-selected communities of data analysis and machine learning professionals.

KDnuggets has had language polls forever. A recent poll is titled “What programming/statistics languages you used for an analytics / data mining / data science work in 2013“. The trends are almost identical to the previous year. The results suggest heavy use of R and Python and SQL for data access. SAS and MATLAB rank higher than I would have expected. I’d expect SAS accounts for larger corporate (Fortune 500) data analysis and MATLAB for engineering, research and student use.

Kaggle offer machine learning competitions and have polled their user base as to the tools and programming languages used by participants in competitions. They posted results in 2011 titledKagglers’ Favorite Tools (also see the forum discussion). The results suggested the abundant use of R. The results also show good use of MATLAB and SAS with much lower Python representation. I can attest that I prefer R over Python for competition work. It just feels though it has more on offer in terms of data analysis and algorithm selection.

Ben Hamner, Kaggle Admin and author of the blog post above on the Kaggle blog goes into more detail on the options when it comes to programming languages for machine learning in a forum post titled “What tools do people generally use to solve problems“.

Ben comments that MATLAB/Octave is a good language for matrix operations and can be good when working with a well defined feature matrix. Python is fragmented by comprehensive and can be very slow unless you drop into C. He prefers Python when not working with a well defined feature matrix and uses Pandas and NLTK. Ben comments that “As a general rule, if it’s found to be interesting for statisticians, it’s been implemented in R” (well said). He also complains about the language itself being ugly and painful to work with. Finally, Ben comments on Julia that doesn’t have much to offer in the way of libraries but is his new favorite language. He comments that it has the conciseness of languages like MATLAB and Python with the speed of C.

Anthony Goldbloom, the CEO of Kaggle gave a presentation to the Bay Area R user group in 2011 on the popularity of R in Kaggle competitions titled Predictive modeling competitions: making data science a sport (see the powerpoint slides). The presentation slides give more detail on the use of programming languages and suggest an Other category that is as close to as large as large as the usage of R. It would be nice to have the raw data that was collected (why didn’t they release it to their own data community, seriously!?).

John Langford on his blog Hunch has an excellent article on the properties of a programming language to consider when working with machine learning algorithms titled “Programming Languages for Machine Learning Implementations“. He divides the properties into concerns of speed and the concerns of programability (programming ease). He points to powerful industry standard implementations of algorithms, all in C and comments that he has not used R or MATLAB (the post was written 8 years ago). Take some time and read some of the comments by academics and industry specialists alike. This is a deep and nuanced problem that really comes down to the specifics of the problem you are solving and the environment in which you are solving it.

Machine Learning Languages

I think of programming languages in the context of the machine learning activities I want to perform.

MATLAB/Octave

I think MATLAB is excellent for representing and working with matrices. As such, I think it’s an excellent language or platform to use when climbing into the linear algebra of a given method. I think it’s suited to learning about algorithms both superficially the first time around and deeply when you are trying to figure something out or go deep into the method. For example, it’s popular in university courses for beginners, like Andrew Ng’s Coursera Machine Learning course.

R

R is a workhorse for statistical analysis and by extension machine learning. Much talk is given to the learning curve, I didn’t really see the problem. It is the platform to use to understand and explore your data using statistical methods and graphs. It has a n enormous number of machine learning algorithms, and advanced implementations too written by the developers of the algorithm.

I think you can explore, model and prototype with R. I think it suits one-off projects with an artefact like a set of predictions, report or research paper. For example, it is the most popular platform for machine learning competitors such as Kaggle.

Python

Python if a popular scientific language and a rising star for machine learning. I’d be surprised if it can take the data analysis mantle from R, but matrix handling in NumPy may challenge MATLAB and communication tools like IPython are very attractive and a step into the future of reproducibility.

I think the SciPy stack for machine learning and data analysis can be used for one-off projects (like papers), and frameworks like scikit-learn may be mature enough to be used in production systems.

Java-family/C-family

Implementing a system that uses machine learning is an engineering challenge like any other. You need good design and developed requirements. Machine learning is algorithms, not magic. When it comes to serious production implementations, you need a robust library or you customize an implementation of the algorithm for your needs.

There are robust libraries, for example Java has Weka and Mahout. Also note that the deeper implementations of core algorithms like regression (LIBLINEAR) and SVM (LIBSVM) are written in C and leveraged by Python and other toolkits. I think you are serious you may prototype in R or Python, but you will implement in a heavier language for reasons such as execution speed and system reliability. For example, the backend of BigML is implemented in Clojure.

Other Concerns

Not a Programmer: If you are not a programmer (or not a confident programmer) I recommend playing machine learning via a GUI interface like Weka.

One Language for Research and Ops: You may want to use the same language for prototyping and for production to reduce risk of not effectively transferring the results.

Pet Language: You may have a pet language of favorite language and want to stick to that. You can implement algorithms yourself or leverage libraries. Most languages have some form of machine learning package, however primitive.

Form 1040

2012

(99)

Department of the Treasury—Internal Revenue Service

US Individual Income Tax Return

For the year Jan 1–Dec 31, 2012, or other tax year beginning

OMB No 1545-0074

, 2012, ending

IRS Use Only—Do not write or staple in this space

See separate instructions

, 20

Your first name and initial

Last name

Your social security number

If a joint return, spouse’s first name and initial

Last name

Spouse’s social security number

Apt no

Home address (number and street) If you have a PO box, see instructions

▲

City, town or post office, state, and ZIP code If you have a foreign address, also complete spaces below (see instructions)

Foreign country name

Filing Status

Check only one

box

Exemptions

Presidential Election Campaign

Check here if you, or your spouse if filing

jointly, want $3 to go to this fund Checking

Foreign postal code

a box below will not change your tax or

refund

You

Spouse

Foreign province/state/county

1

4

Single

Married filing jointly (even if only one had income)

2

3

c

Head of household (with qualifying person) (See instructions) If

the qualifying person is a child but not your dependent, enter this

child’s name here ▶

Married filing separately Enter spouse’s SSN above

and full name here ▶

6a

b

5

Qualifying widow(er) with dependent child

Yourself If someone can claim you as a dependent, do not check box 6a

Spouse

Dependents:

(1) First name(2) Dependent’s

social security number

Last name

}

(4) ✓ if child under age 17

qualifying for child tax credit

(see instructions)

(3) Dependent’s

relationship to you

Dependents on 6c

not entered above

d

Attach Form(s)

W-2 here Also

attach Forms

W-2G and

1099-R if tax

was withheld

If you did not

get a W-2,

see instructions

Enclose, but do

not attach, any

payment Also,

please use

Form 1040-V

Adjusted

Gross

Income

Boxes checked

on 6a and 6b

No of children

on 6c who:

• lived with you

• did not live with

you due to divorce

or separation

(see instructions)

If more than four

dependents, see

instructions and

check here ▶

Income

Make sure the SSN(s) above

and on line 6c are correct

8b

8a9a

10

11

Qualified dividends

9b

Taxable refunds, credits, or offsets of state and local income taxes

Alimony received

10

11

12

13

14

Business income or (loss) Attach Schedule C or C-EZ

Capital gain or (loss) Attach Schedule D if required If not required, check here ▶

Other gains or (losses) Attach Form 4797

12

13

14

15a

16a

17

IRA distributions

15a

b Taxable amount

Pensions and annuities 16a

b Taxable amount

Rental real estate, royalties, partnerships, S corporations, trusts, etc Attach Schedule E

15b

16b

17

18

19

20a

Farm income or (loss) Attach Schedule F

Unemployment compensation

Social security benefits 20a

18

19

20b

21

22

Other income List type and amount

Combine the amounts in the far right column for lines 7 through 21 This is your total income

23

Educator expenses

24

Certain business expenses of reservists, performing artists, and

fee-basis government officials Attach Form 2106 or 2106-EZ

25

Health savings account deduction Attach Form 8889

24

25

26

27

28

Moving expenses Attach Form 3903

Deductible part of self-employment tax Attach Schedule SE

Self-employed SEP, SIMPLE, and qualified plans

26

27

28

29

30

31a

Self-employed health insurance deduction

Penalty on early withdrawal of savings 32

33

34

Alimony paid b Recipient’s SSN ▶

IRA deduction

Student loan interest deduction

Tuition and fees Attach Form 8917

29

30

31a

32

33

34

35

36

37

Domestic production activities deduction Attach Form 8903

35

Add lines 23 through 35

Subtract line 36 from line 22 This is your adjusted gross income

7

8a

b

9a

b

Total number of exemptions claimed

Wages, salaries, tips, etc Attach Form(s) W-2

Taxable interest Attach Schedule B if required

Tax-exempt interest Do not include on line 8a

Ordinary dividends Attach Schedule B if required

b Taxable amount

▶Add numbers on

lines above ▶

7

21

22

23

For Disclosure, Privacy Act, and Paperwork Reduction Act Notice, see separate instructions

▶

36

37

Cat No 11320B

Form

1040

(2012)

Page 2

Form 1040 (2012)

Tax and

Credits

Standard

Deduction

for—

• People who

check any

box on line

39a or 39b or

who can be

claimed as a

dependent,

see

instructions

• All others:

Single or

Married filing

separately,

$5,950

Married filing

jointly or

Qualifying

widow(er),

$11,900

Head of

household,

$8,700

Other

Taxes

Payments

If you have a

qualifying

child, attach

Schedule EIC

38

Amount from line 37 (adjusted gross income)

39a

Check

if:

Sign

Here

Paid

Preparer

Use Only

You were born before January 2, 1948,

Spouse was born before January 2, 1948,Blind

Blind

}

38

Total boxes

checked ▶ 39a

39b

42

43

Exemptions Multiply $3,800 by the number on line 6d

Taxable income Subtract line 42 from line 41 If line 42 is more than line 41, enter -0-

Form 4972 c

962 election

Tax (see instructions) Check if any from: a

Form(s) 8814 b

44

45

46

Alternative minimum tax (see instructions) Attach Form 6251

Add lines 44 and 45

b

8919Foreign tax credit Attach Form 1116 if required

Credit for child and dependent care expenses Attach Form 2441

47

48

49

50

51

Education credits from Form 8863, line 19

Retirement savings contributions credit Attach Form 8880

Child tax credit Attach Schedule 8812, if required

49

50

51

52

53

54

55

Residential energy credits Attach Form 5695

52

3800 b

8801 c

Other credits from Form: a

53

Add lines 47 through 53 These are your total credits

Subtract line 54 from line 46 If line 54 is more than line 46, enter -0-

56

57

Self-employment tax Attach Schedule SE

Unreported social security and Medicare tax from Form:

58

59a

b

Additional tax on IRAs, other qualified retirement plans, etc Attach Form 5329 if required

Household employment taxes from Schedule H

a

4137

▶

First-time homebuyer credit repayment Attach Form 5405 if required▶

69

Credit for federal tax on fuels Attach Form 4136

70

Credits from Form: a

2439 b

Reserved c

8801 d

8885 71

Add lines 62, 63, 64a, and 65 through 71 These are your total payments

▶Other taxes Enter code(s) from instructions

62

63

Federal income tax withheld from Forms W-2 and 1099

2012 estimated tax payments and amount applied from 2011 return

64a

b

Earned income credit (EIC)

Nontaxable combat pay election

64b

Additional child tax credit Attach Schedule 8812

Add lines 55 through 60 This is your total tax

American opportunity credit from Form 8863, line 8

Reserved

Amount paid with request for extension to file 44

45

46

55

56

57

60

61

62

63

64a

65

66

67

68

Excess social security and tier 1 RRTA tax withheld

73

If line 72 is more than line 61, subtract line 61 from line 72 This is the amount you overpaid

74a

b

d

Amount of line 73 you want refunded to you If Form 8888 is attached, check here ▶

▶ c Type:

Routing number

Checking

Savings

Account number

Amount of line 73 you want applied to your 2013 estimated tax ▶ 75

Amount you owe Subtract line 72 from line 61 For details on how to pay, see instructions ▶

75

76

43

58

59a

59b40

41

42

54

▶60

61

65

6647

48

77

77

Estimated tax penalty (see instructions)

Do you want to allow another person to discuss this return with the IRS (see instructions)?

Designee’s

name ▶

Phone

no ▶

72

73

74a

76

Yes Complete below

No

Personal identification

▶

number (PIN)

Under penalties of perjury, I declare that I have examined this return and accompanying schedules and statements, and to the best of my knowledge and belief,

they are true, correct, and complete Declaration of preparer (other than taxpayer) is based on all information of which preparer has any knowledge

Your signature

Date

Your occupation

Daytime phone number

Spouse’s signature If a joint return, both must sign

Date

Spouse’s occupation

If the IRS sent you an Identity Protection

PIN, enter it

here (see inst)

PTIN

Check

if

self-employed

▲

Joint return? See

instructions

Keep a copy for

your records

Itemized deductions (from Schedule A) or your standard deduction (see left margin)

Subtract line 40 from line 38

▶

Third Party

Designee

40

41

Direct deposit?

See

▶

instructions

Amount

You Owe

If your spouse itemizes on a separate return or you were a dual-status alien, check here ▶

b

67

68

69

70

71

72

Refund

{

Print/Type preparer’s name

Firm’s name

Preparer’s signature

Date

▶

Firm's EIN

Firm’s address ▶

Phone no

▶

Form 1040 (2012)