

基于多种机器学习算法的游客体验 评价分析系统

V1.0

用户手册

目录

1 引言	4
1.1 编写目的	4
1.2 参考资料	4
1.3 术语和算法简要解释	6
2 软件概述	6
2.1 软件功能	8
2.2 软件运行	8
2.3 系统要求	8
3 系统使用	10
3.1 运行准备	10
3.1.1 设置默认变量	
3.1.2 提供预测文本	
3.1.3 提供训练文本（仅当需要训练模型时）	
3.2 运行系统	Error! Bookmark not defined.
3.2.1 方法一：使用 Atom 运行系统	
3.2.2 方法二：使用命令提示符运行系统	
3.3 获取分析报告	Error! Bookmark not defined.
3.3.1 方法一：使用 Atom 修改文字过滤器	
3.3.2 方法二：使用命令提示符修改文字过滤器	
3.4 修改文字过滤器	Error! Bookmark not defined.
3.4.1 方法一：使用 Atom 运行并修改系统	
3.4.2 方法二：使用命令提示符运行并修改系统	
3.5 训练模型并修改模型	Error! Bookmark not defined.
3.5.1 方法一：使用 Atom 运行并修改模型	
3.5.2 方法二：使用命令提示符运行并修改模型	

1 引言

1.1 编写目的

很多网站每天都会收到海量的用户反馈，通过人力来筛分评级是极为困难且不经济的。这其中，旅游业的用户评价多元而复杂，而潜在用户在选择线路和旅游公司时却更加依赖之前客户评价。

目前已有的游客体验评价系统主要存在以下问题（1）用户往往从多个方面点评，传统线性判别分析准确率低（2）深度神经网络训练时间过长，并且由于用户评价篇幅往往长短相差悬殊，表达清晰和模糊有别，导致训练模型只对一类评价过度拟合，在分析大量评价时准确率不高。（3）情感分析的词库多采用默认的单词库，例如微软的 `nltk` 词库或者 `Corpus` 的词库，但不同公司会有不同的需要重点分析的词汇，需要让用户可以自行增加词库内容（4）市面上不同机器学习模型有不同的适用范围，需要有某个系统能够实现多种模型的拟合，并选取最佳机器学习模型。

1.2 参考资料

技术规范：

《软件需求规格说明书》

参考文献：

- [1] Emir, A., Halim, H., Hedre, A., Abdullah, D., Azmi, A., & Kamal, S. B. M. (2016). Factors influencing online hotel booking intention: A conceptual framework from stimulus-organism-response perspective. *International Academic Research Journal of Business and Technology*, 2(2), 129-134.
- [2] Kim, W. G., & Park, S. A. (2017). Social media review rating versus traditional customer satisfaction: which one has more incremental predictive power in explaining hotel performance?. *International Journal of Contemporary Hospitality Management*, 29(2), 784-802.
- [3] Jianqiang, Z., Xiaolin, G., & Xuejun, Z. (2018). Deep convolution neural networks for twitter sentiment analysis. *IEEE Access*, 6, 23253-23260.
- [4] data.world (2019) Hotel Reviews - dataset in DATAFINITY. <https://data.world/datafiniti/hotel-reviews>.
- [5] Hao, J., & Ho, T. K. (2019). Machine Learning Made Easy: A Review of Scikit-learn Package in Python Programming Language. *Journal of Educational and Behavioral Statistics*, 44(3), 348-361.

1.3 术语和算法简要解释

机器学习 (Machine Learning)：一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。

逻辑回归算法(logistic regression)：一种传统线性拟合算法，逻辑回归通过使用其固有的逻辑函数估计概率，来衡量因变量（我们想要预测的标签）与一个或多个自变量（特征）之间的关系。Logistic 回归是二分类任务的首选方法。它输出一个 0 到 1 之间的离散二值结果。简单来说，它的结果不是 1 就是 0。

决策树算法 (decision tree)：一种机器学习算法，决策树算法是随机森林算法的基础。信息论中有熵 (entropy) 的概念，表示状态的混乱程度，熵越大越混乱。熵的变化可以看做是信息增益，决策树算法的核心思想是以信息增益度量属性选择，选择分裂后信息增益最大的属性进行分裂，在每次分裂的时候贪心选择信息增益最大的属性，作为本次分裂属性。每次分裂就会使得树长高一层。这样逐步生产下去，就一定可以构建一颗决策树。

随机森林算法 (random forest)：一种机器学习算法。通过训练多个决策树算法模型，生成的多棵决策树组成随机森林。对于分类问题，按多棵树分类器投票决定最终分类结果；对于回归问题，由多棵树预测值的均值决定最终预测结果

支持向量机 (support vector machine): 一种机器学习算法。支持向量机是一种二分类模型,它的基本模型是定义在特征空间上的间隔最大的线性分类器,模型的基本想法是求解能够正确划分训练数据集并且几何间隔最大的分离超平面。

朴素贝叶斯(Naïve Bayesian): 朴素贝叶斯分类器是一系列以假设特征之间强(朴素)独立下运用贝叶斯定理为基础的简单概率分类器。朴素贝叶斯分类器是高度可扩展的,因此需要数量与学习问题中的变量(特征/预测器)成线性关系的参数。最大似然训练可以通过评估一个封闭形式的表达式来完成, [1]:718 只需花费线性时间,而不需要其他很多类型的分类器所使用的费时的迭代逼近。

神经网络(neural network): 一种机器学习算法,是一种模仿生物神经网络(动物的中枢神经系统,特别是大脑)的结构和功能的数学模型或计算模型,用于对函数进行估计或近似。神经网络由大量的人工神经元联结进行计算。大多数情况下人工神经网络能在外界信息的基础上改变内部结构,是一种自适应系统。

自然语言库(Natural Language Toolkit): 由微软公司发布,简称 nltk,包含大量语言中使用中性词汇,例如 he, she, who 等,可用于筛选句子段落中无意义的词汇,从而提高模型预测效果。

2 软件概述

2.1 软件功能

本软件著作权专利主要可以做到：

(1) 当用户提供在 excel 文件中的文本文字后（如果要处理多行文字，则放置在 excel 的不同行中），运行本系统，系统会输出计算结果，0 表示情感消极，1 表示情感积极。经测试，对于 10000 行平均约 100 字节的用户文字数据，本软件情感分析预测用时不超过 5 秒

(2) 用户可根据公司实际处理的用户反馈文字，除使用默认的微软 nltk 词库外，增加情感分析词汇。

(3) 用户可以提供已经标签过的数据，也就是在 excel 文件中第一列为不同的客户的反馈文字，第二列为 0 或者 1。然后系统会进行训练，输出和给定 0-1 序列最接近的预测模型，该模型可用于客户之后的模型训练。经测试，对于 10000 行平均约 100 字节的用户文字以及标签过 0 或者 1 的数据，本软件模型训练用时不超过 2 分钟。

2.2 软件运行

本系统基于 Python3 语言编写，可运行在 PC 及其兼容机上，使用 WINDOWS 或者 LINUX（例如 Ubuntu 等）操作系统。

程序主要有以下两种运行方法，详细操作和截图请见 3.2。

方法一：使用 Pycharm, Atom 等免费应用软件运行 Python。

方法二：使用命令提示符应用（command prompt APP）运行 Python。

2.3 系统要求

Windows XP 以上系统，256M 以上内存。

系统已安装 Python3

如果不了解命令提示符应用，则需要至少已经安装 Pycharm, Atom 等可运行 Python 代码的某一个免费应用软件。

2.4 系统流程图

系统流程如图 2-1 所示。

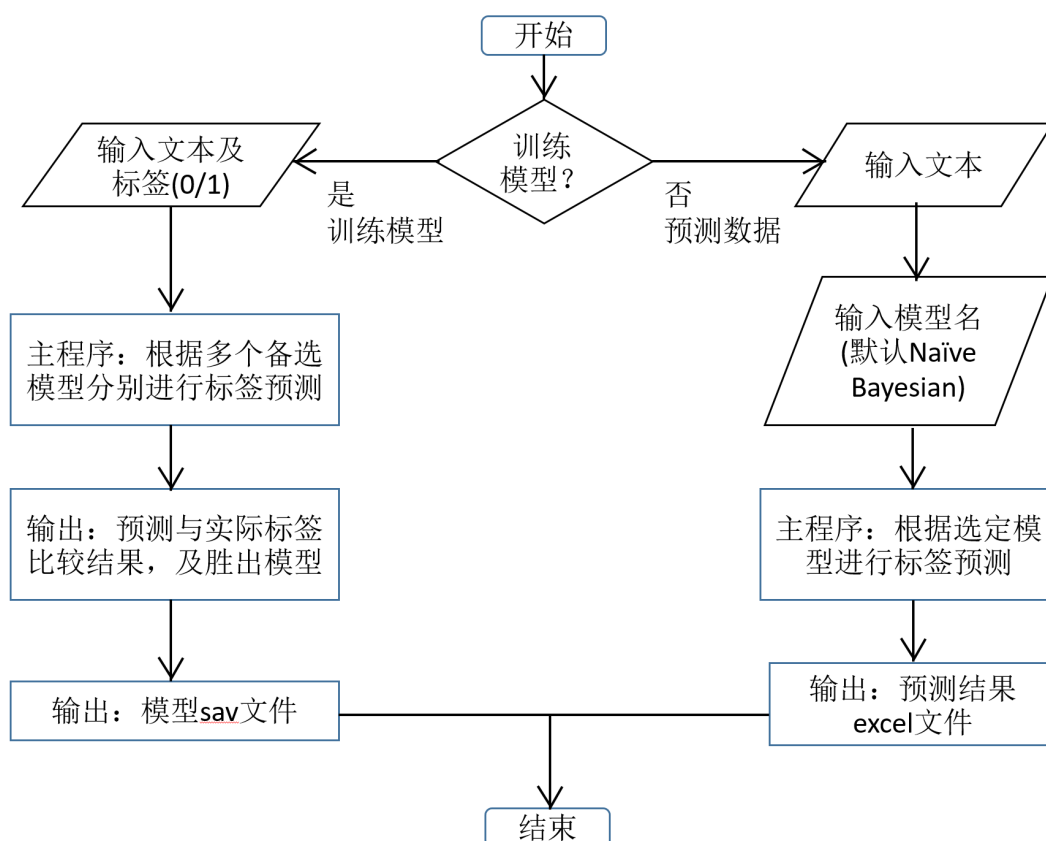


图 2-1 系统流程图

3 系统使用

3.1 运行准备

本系统默认读取与“Xiaoyu_Li_software_coding.py”文件在同一文件夹中的文件，各文件名及功能及功能见表 3-1。

表 3-1 文件名及功能

文件名	功能
Xiaoyu_Li_software_coding.py	代码部分，可运行
dictionary.xlsx	模型训练及预测的词库，所有运算基于该词库，不属于该词库的单词不会进行计算。
predict_text.xlsx	用户提供需要预测的文本于该 excel 文件中列名称为' text' 的列
train_text.xlsx	用户提供需要训练的文本于该 excel 文件中列名称为' text' 的列，并提供 0 或者 1 的标签于列名称为' label' 的列
Models/ 文件夹	包含多个 sav 为后缀的模型文件，这些文件为已经训练完成的文件，可修改图 3-X 或者图 3-XX 中

	的 model_name 变量名后直接调用。
--	------------------------

所有已训练完成的模型在和“Xiaoyu_Li_software_coding.py”文件在同一目录内的 mode/文件夹中，在表 3-2 中列出了文件名所对应的各个机器学习模型名称。关于模型的介绍请见说明书 1.3 节。

表 3-2 模型文件名与机器学习模型

文件名	机器学习模型
code_model_DecisionTreeClassifier.sav	决策树模型
code_model_LogisticRegression.sav	逻辑回归模型
code_model_MLPClassifier.sav	神经网络模型
code_model_MultinomialNB.sav	朴素贝叶斯模型
code_model_RandomForestClassifier.sav	随机森林模型
code_model_SVC.sav	线性支持向量机模型

3.2 运行系统

3.2.1 方法一：使用 Atom 运行系统

Atom 是一个使用简便的免费的代码运行程序，可支持 Window 和 Linux 操作系统。该软件可在该链接下载：

<https://atom.io/>

打开上述链接，如图 3-1，网站可自动监测电脑的系统，点击”download”下载安装包，安装过程非常简单。

安装完成后，打开 Atom 软件，在软件的欢迎界面点击 “Install a Package”，如果 3-2 所示。



图 3-1 Atom 下载界面

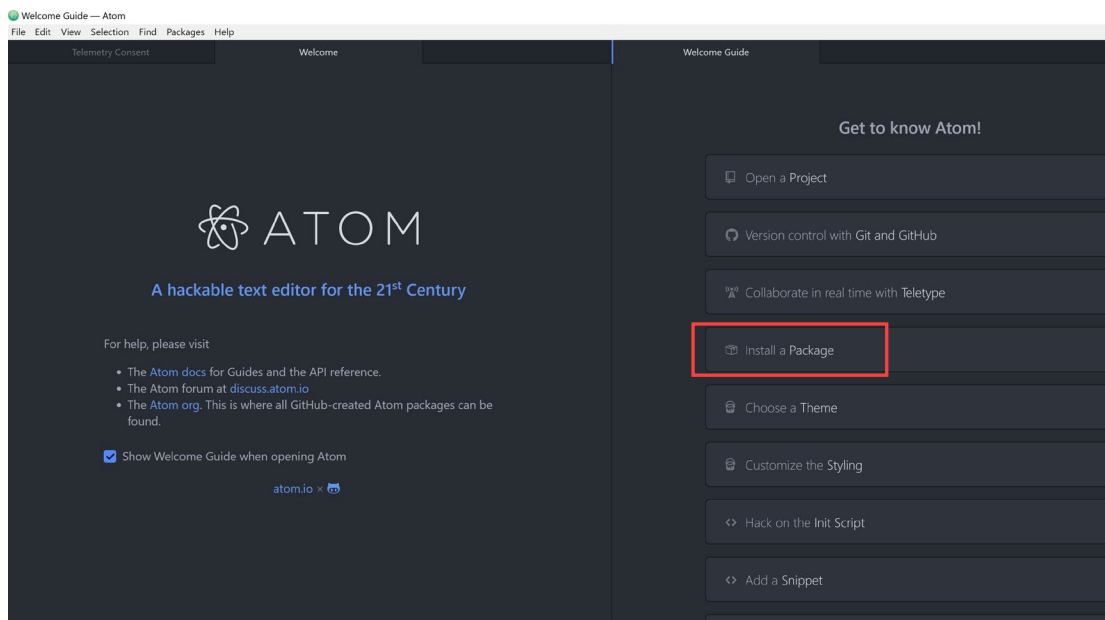


图 3-2 Atom 欢迎界面

如图 3-3 所示，接下来点击 “Open installer”，则会弹出图 3-4 所示的” Install Packages” 对话框。 在对话框中输入 “script”，

选择第一个返回的搜索内容，点击“Install”，约一分钟后完成安装。

注意，由于作者已经安装过“script”包，图 3-4 中显示的是 Uninstall（卸载），对于第一次使用的客户，应显示与“server-script”类似的”Install”图标。

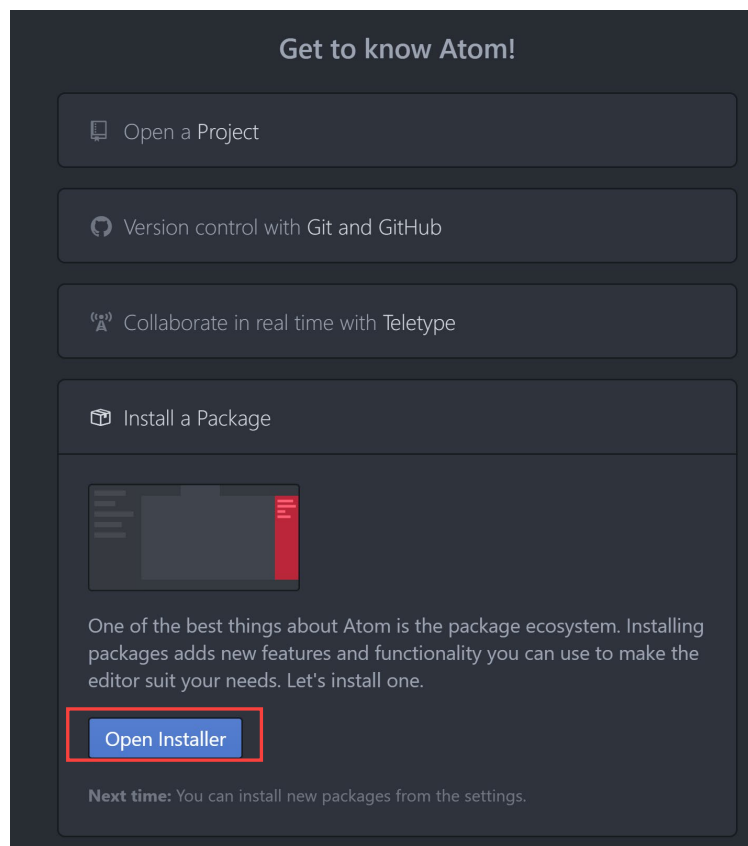


图 3-3 点击“Install a Package”

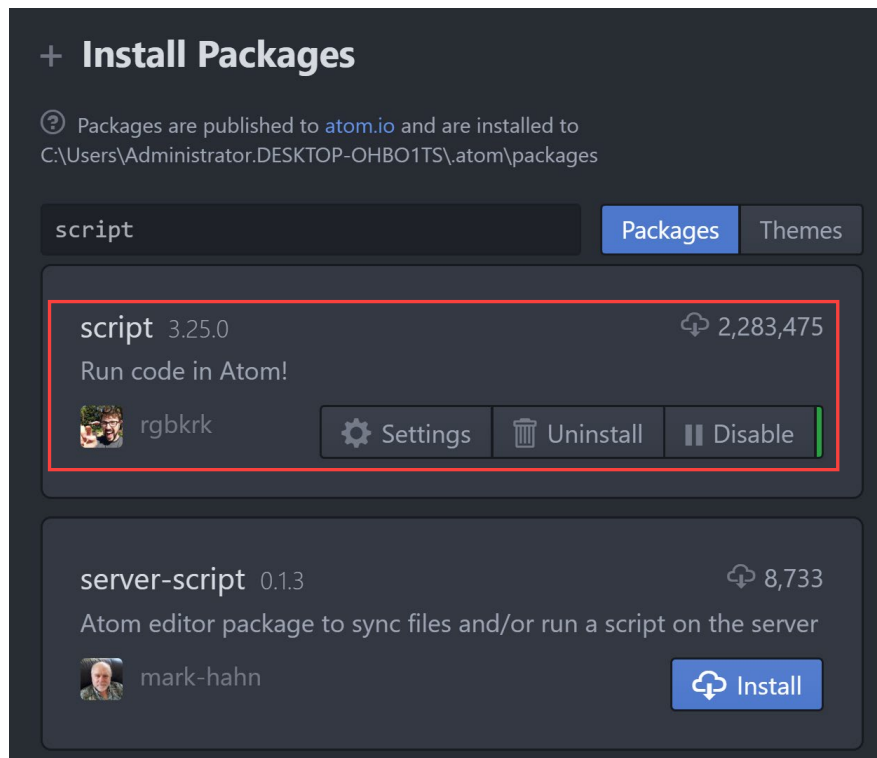


图 3-4 下载“script”包

接下来点击 Atom 上方的工具栏的“file”，接下来点击“Open file”，如图 3-5（a）所示，然后点击文件“Xiaoyu_Li_software_coding.py”。如图 3-5（b）所示，代码在 Atom 中打开。

特别注意，如图 3-6 所以为用户可修改部分，一般来说用户不需要修改。如果需要修改，具体操作在 3.4.1 与 3.5.1 节介绍。

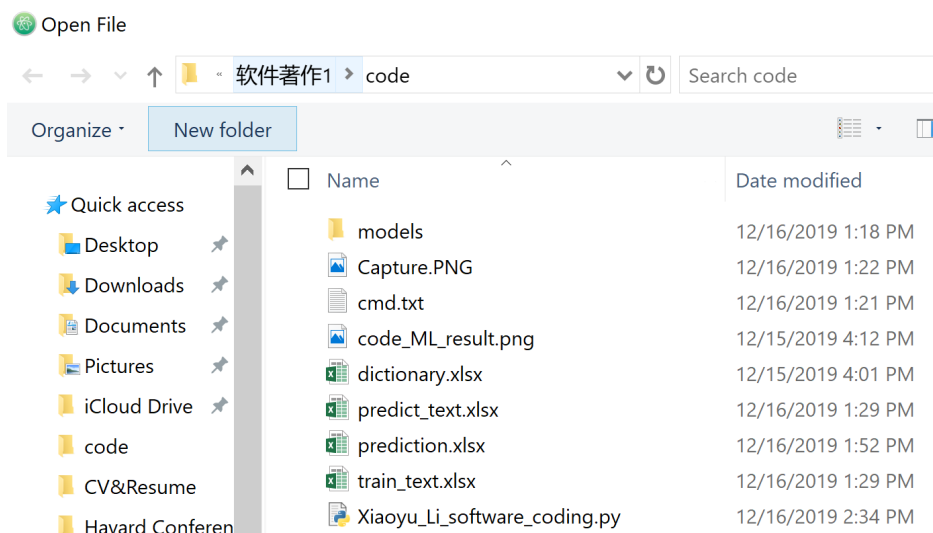


图 3-5 (a) 使用 Atom 打开代码

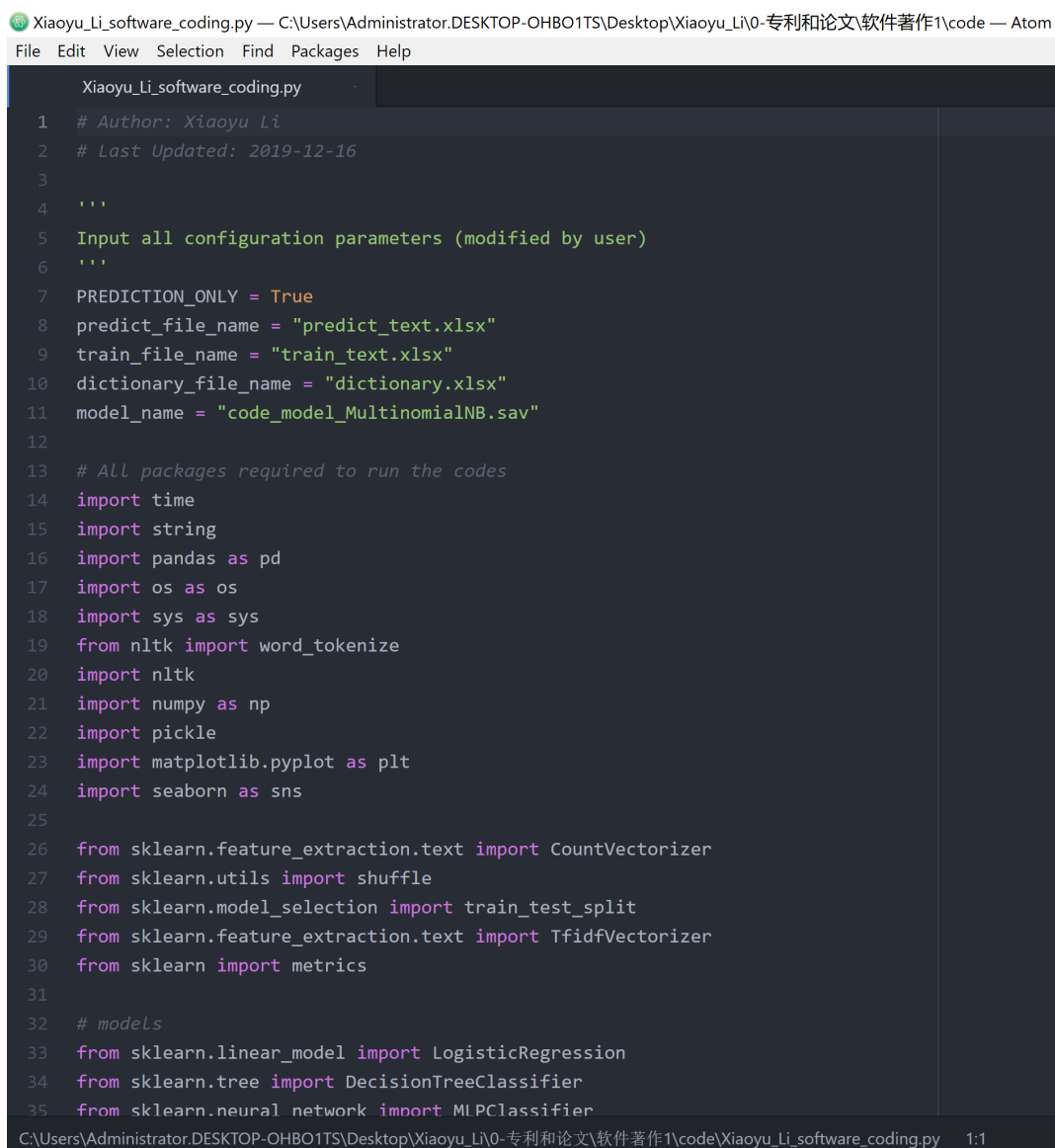


图 3-5 (b) 使用 Atom 打开代码

```
# Author: Xiaoyu Li
# Last Updated: 2019-12-16

'''
Input all configuration parameters (modified by user)
'''

PREDICTION_ONLY = True
predict_file_name = "predict_text.xlsx"
train_file_name = "train_text.xlsx"
dictionary_file_name = "dictionary.xlsx"
model_name = "code_model_MultinomialNB.sav"
```

图 3-6 用户可修改部分(通过 Atom 软件打开)

如图 3-7 所示,需要预测的文本输入在 predict_text.xlsx 的“text”列中。其它列如果有数据不影响,本软件只根据“text”为关键词读入该列的所有行的文字用于预测。

如图 3-8 所示,在 Atom 中选择 Packages→Script→Run Script 或者同时按键盘的”CTRL+Shift+B”,则 Atom 开始运行 Python 文件。

如图 3-9 所示, prediction.xlsx 文件被生成。具体解释见 3.3 节。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	text																			
2	This hotel was nice and quiet. Did not know, there was train track near by. But it was only few train passed during our stay. Best Western changed hotel classification. The Plus category are not the same as before.																			
3	We stayed in the king suite with the separation between the bedroom and the living space. The sofa bed wasn't very good I had back discomfort by the day we left on our three night stay. The room is clean, and the king bed very comfy																			
4	Parking was horrible, somebody ran into my rental car while staying there. I didn't get to try the breakfast, I was there for business so the restaurant opened to late for the business world to enjoy. I had to asked for coffee for a																			
5	Not cheap but excellent location. Price is somewhat standard for not having reservations. But room was nice and clean. They offer good continental breakfast which is a plus and compensates. Front desk service and personnel where exc																			
6	If you get the room that they advertised on the website and for what you paid, you may be lucky. If you stay many days, they will give you the not so good rooms. Nobody wants to stay in these rooms: low light/dark rooms, near pool, a																			
7	This is such a fun, lovely hotel. The attention to detail is impressive. from the thicker rimmed water glasses to the extra fluffy towels. Loved the vibrant art which lends itself to a hip vibe. My only disappointment was at their r																			
8	We recently stayed at this hotel on a trip to Lexington with other friends. Our group shared the feeling that we would not be back. We routinely waited more than 10 minutes for elevators. The workout room is tiny with 2 treadmills a																			
9	I reserved a room a week in advance, knowing a hotel is usually not great accommodations but we were just passing thru during the longest yard sale. I was quoted 50 over the phone and given our room numbers as a confirmation number																			
10	Wowhere																			
11	Hello, I have traveled a lot and abroad and by far this is the worst place i have ever booked. i paid got the key and walked in. OMG the place is HORRIBLE. this place is one of those hotels they show in horror movies where they kill																			
12	This is a new Hampton Inn and is well decorated. The color scheme gives a feeling of luxury and is very restful. The beds are comfortable and the black out curtains keep the room dark for a good nights sleep. The entrance is a littl																			
13	I stayed here for three nights while I explored nearby Santa Barbara. This place is ok for a night or two. if you can live the excess noise coming from the freeway, then three nights is just about doable																			
14	Completely accommodating hotel! I booked a reservation last minute day of, used points to offset the cost and even called the front desk to ask for a little help. They were completely kind and welcoming, especially accommodating whi																			
15	The water is very hot and there's no cold water at all... It's awful. The food is not good. The staff is friendly and very nice. More upscale items should be offered for the price of the room at 200. A night! Snacks, toiletries,																			
16	Great staff and rooms. Housekeeping was always spot on—we never saw them come in, and it was in perfect order when we came back to our rooms. The hotel is close to everything, French Quarter, museums, business district, etc. And we																			
17	This hotel, formerly a prestigious bank, may be attractive for history buffs, but is an expensive and horrible place to stay when you visit New Orleans. The rooms are SMALL, unattractive, with terrible old bedding, virtually without																			
18	Very accommodating staff. Competitive pricing for downtown N.O. convenient location. Comfortable room. Good food in restaurant. Broken exercise equipment. 'fitness club' is a bit of a misnomer. We can recommend this to anyone else f																			
19	Room was much larger than I expected, and water and fridge were offered for my room - had everything I needed. Nice basic breakfast was included. Room was very clean and everyone here was friendly! Would stay again. We're thrilled t																			
20	The Whitney Hotel is ideally located to see most of the sights of New Orleans. I would say it is a place to lay your head but I found the restaurant poor in terms of food choice and quality. The room phone did not work, no room serv																			
21	This is an old hotel in very good location. A 5 minute walk to Bourbon Street and Barracks' Casino. A 15 min walk will get you to Gaffe, Oldside and the French Market. Although old, the hotel is well kept and the staff is efficient. I																			
22	The hotel is historic and yet modern. Full of history. The hotel is conveniently located in the heart of New Orleans. Blocks from the French Quarter and the Riverfront. The staff was very courtis and very knowledgeable. Always there																			
23	This hotel aptly named the Whitney Hotel as it was the location of the Whitney Bank. Several of the features of the bank era are there to be seen such as the strong room doors and related features. This hotel is in a very good locatio																			
24	Such a beautiful hotel! LOVED the lobby and the bank vault. It was so unique and interesting. Our suite was so beautiful and very comfortable. We would like to thank the staff for taking such good care of us. Your knowledge of your v																			
25	We recently had the opportunity to travel to New Orleans and we chose to stay at the Whitney based on user reviews. Location, and price. The room was very nice and was really clean. The hotel was well decorated. The one big call out																			
26	Our room was beautiful, the location was wonderful. Just a couple of blocks from shopping, Bourbon Street and an easy walk to Jackson Square. There are great restaurants in walking distance as well as the restaurant in the hotel. Din																			
27	Excellent service and location was perfect. Have been in New Orleans numerous times so we didnt need to be in the heart of French Quarter which could be very noisy, dont get me wrong, love a good time let the good times roll but whe																			
28	I will never consider staying at the Whitney after an epic night of uncomfortable room temperatures and a night's sleep on the hardest, cheapest mattress I've ever experienced. The hotel staff was nonexistent and are complaints wer																			
29	The hotel was alright, this was my second stay at this hotel after my first experience I recieved a comp night to get a better experience. My second experience was about the same, rude staff, bugs in room, room wasn't cleaned when a																			
30	Nice location! Good staff. Th rooms need to be Aired and to be cleaned. Th breakfast is a shame! Th location is good! I dont recommend this hotel as for th quality of th rooms as for the quality of th breakfast server into plastic plates																			
31	Recently stayed at the Whitney Hotel in New Orleans. The hotel was an exceptional experience from the old 1919 building to the amazing queen suite. We would make any future visits to New Orleans include a stay at the Whitney Hotel.																			
32	Unfortunately, it is not a place to have confidence. One of my children left a small amount of money in the car. When we left the hotel we realized they had stolen it. Although it was little amount of money, the fact is equally rep																			
33	My husband and I stayed for 4 nights with my 13 month old son and our experience was disappointing from the moment we checked in. When we made our reservation on the phone, we requested a crib for my son and were told it would be no																			
34	What a wonderful weekend spent with my husband. Staff, Food, Rooms, everything was great. So welcoming and professional from arriving to departure. Location is perfect. Not too busy, but many thing within walking distance to explore																			
35	This was a beautiful hotel. All the things that made it a bank are still there, reminding you of the lovely banks that we used to have. The fact that part of the building is still an operating bank was a surprise. Everything was lov																			

图 3-7 读入 predict_text.xlsx 文件

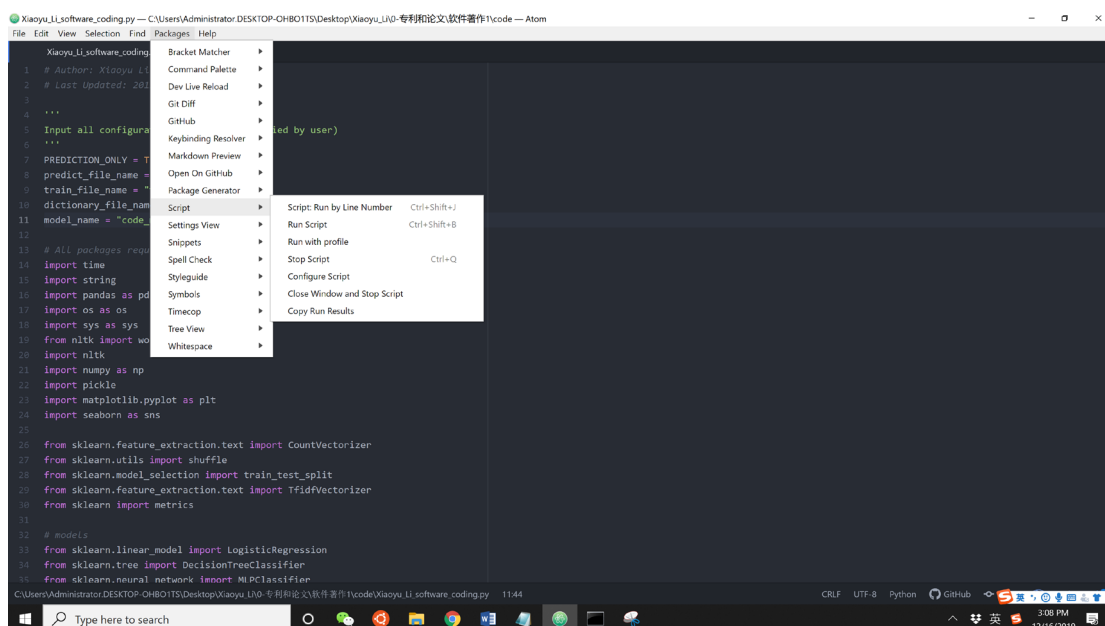


图 3-8 运行本软件(通过 Atom 软件)

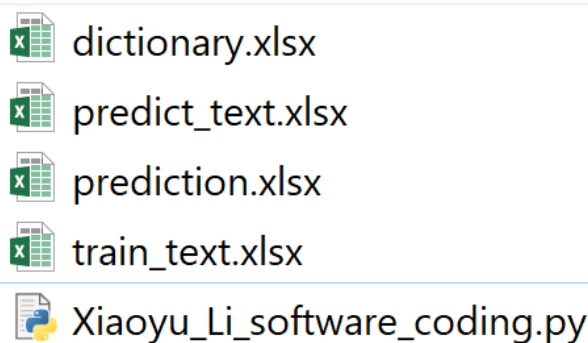


图 3-9 运行代码并生成 prediction.xlsx 文件

3.2.2 方法二：使用命令提示符运行系统

注意，本方法适用于有一定计算机系统了解的用户。用户需要使用 Window shell 或者 Linux shell 命令的知识。

作者使用的是 Linux 系统如图 3-10，在“搜索”中输入 cmd，然后按键盘“回车”键，则命令提示符(command prompt)被打开。

通过 cd 命令让命令提示符进入

到“Xiaoyu_Li_software_coding.py”所在的目录。关于 cd 命令的教程可见：

<https://www.runoob.com/linux/linux-comm-cd.html>

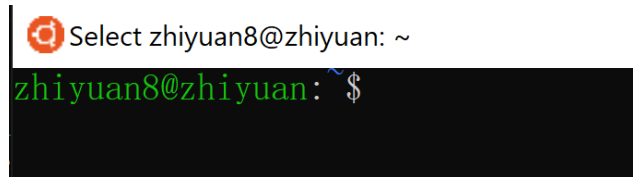


图 3-10 打开命令提示符

如图 3-11 所示，当进入到” Xiaoyu_Li_software_coding.py” 所在的目录后，输入下列指令运行 Python 文件：

```
python3 Xiaoyu_Li_software_coding.py
```

运行会输出 predict.xlsx 文件，如图 3-9 所示。程序的中间输出结果如图 3-12 所示。

```
current working directory is : /mnt/c/Users/Administrator.DESKTOP-OHBOITS/Desktop/Xiaoyu_Li/0-专利和论文/软件著作1/code
**** Read in 20158 number of words as dictionary ****
**** Read in 7086411 bytes Customer text****
load_model: MultinomialNB(alpha=1.0, class_prior=None, fit_prior=True)
===== a head view of predictions =====
      text  Prediction
0  This hotel was nice and quiet. Did not know, t...      1
1  We stayed in the king suite with the separatio...      1
2  Parking was horrible, somebody ran into my ren...      0
3  Not cheap but excellent location. Price is som...      0
4  If you get the room that they advertised on th...      0
****finished saving prediction to /mnt/c/Users/Administrator.DESKTOP-OHBOITS/Desktop/Xiaoyu_Li/0-专利和论文/软件著作1/code/prediction.xlsx****
```

图 3-11 运行本软件(通过命令提示符)

3.3 获取分析结果

根据 prediction.xlsx 预测的结果见图 3-12。

例如 predict_text 的第一行文字为：

This hotel was nice and quiet. Did not know, there was train track near by. But it was only few train passed during our stay. Best Western changed hotel classification. The Plus category are not the same as before.

我们知道这个用户的总体评价是偏好的，提到了这个宾馆很好(nice)而且安静(quiet)。因而图 3-10 中预测结果是 1，表示积极。

又例如 predict_text 的第 3 行文字为：

Parking was horrible, somebody ran into my rental car while staying there. I didn't get to try the breakfast, I was there for business so the restaurant opened to late for the business world to enjoy, I had to asked for coffee for my room, And the items in the vending machine were stale.

我们知道这个用户的总体评价是偏糟糕的，提到了这个宾馆停车糟糕 (horrible) 而且设施陈旧 (stale)。因而图 3-10 中预测结果是 0，表示消极。

	A	B	C
1		text	Prediction
2	0	This hotel was nice and quiet. Did not know, t	1
3	1	We stayed in the king suite with the separatic	1
4	2	Parking was horrible, somebody ran into my rer	0
5	3	Not cheap but excellent location. Price is son	0
6	4	If you get the room that they advertised on th	0
7	5	This is such a fun, lovely hotel. The attentio	1

图 3-12 打开 prediction.xlsx 文件

3.4 修改文字过滤器

文字分析的词库依据 dictionary.xlsx 文件，内含约 2 万个单词。用户可以自行删除或者增加词汇。注意，只有英语或者数字可以读入，目前本系统尚不支持中文。

如果需要用在中文的客户反馈情感分析，那么需要将文字通过百度，谷歌等文本翻译助手翻译成英文后输入本系统。

1	dictionary
2	abandon
3	abandoned
4	abbey
5	abbiamo
6	abbondante
7	abby
8	abiding
9	abilities
10	ability
11	abit
12	able
13	ables
14	ablesgeneral
15	ablt

图 3-13 打开 dictionary.xlsx 文件

3.5 训练模型并修改模型

3.5.1 方法一：使用 Atom 运行系统

打开 Atom,如图 3-14 所示,将 PREDICTION_ONLY 由 True 修改为 False,注意第一个字母大写。

训练模型需要提供训练文本和文本标签。如图 3-15 所示,在 train_text.xlsx 的“text”列输入客户的反馈文字。比如宾馆获得的是 1-5 星的评价,由于标签只能是 0 或者 1,通过 EXCEL 的 IF-ELSE 基本命令:

=IF(B2>=4, 1, 0)

将 1-3 星评价定义为“0”(消极),4-5 星评价定义为“1”(积极)。按照 3-4 节方法运行 Python 代码。

```
'''
Input all configuration parameters (modified by user)
'''
PREDICTION_ONLY = False
predict_file_name = "predict_text.xlsx"
train_file_name = "train_text.xlsx"
dictionary_file_name = "dictionary.xlsx"
model_name = "code_model_MultinomialNB.sav"
```

图 3-14 修改 PREDICTION_ONLY 变量

1	text	rating	label
2	This hotel was nice and quiet. Did not know, there was trai	3	0
3	We stayed in the king suite with the separation between the	4	1
4	Parking was horrible, somebody ran into my rental car while	3	0
5	Not cheap but excellent location. Price is somewhat standar	5	1
6	If you get the room that they advertised on the website and	2	0

图 3-15 输入 train_text.xlsx 内容







-  code_model_DecisionTreeClassifier.sav
-  code_model_LogisticRegression.sav
-  code_model_MLPClassifier.sav
-  code_model_MultinomialNB.sav
-  code_model_RandomForestClassifier.sav
-  code_model_SVC.sav

图 3-16 训练获得的模型

运行代码后训练得到的模型将存储于 model/文件夹中。此外，模型训练比较存储为图片文件 code_ML_result.png。结果如图 3-17 所示，简单起见一般只需要关注 Prediction Accuracy（预测准确率）。从图 3-17 中可知，朴素贝叶斯模型 (Naïve Bayesian) 预测效果最好。如图 3-6，下次再预测模型时可将 “model_name” 变量名修改为 “code_model_MultinomialNB.sav”。

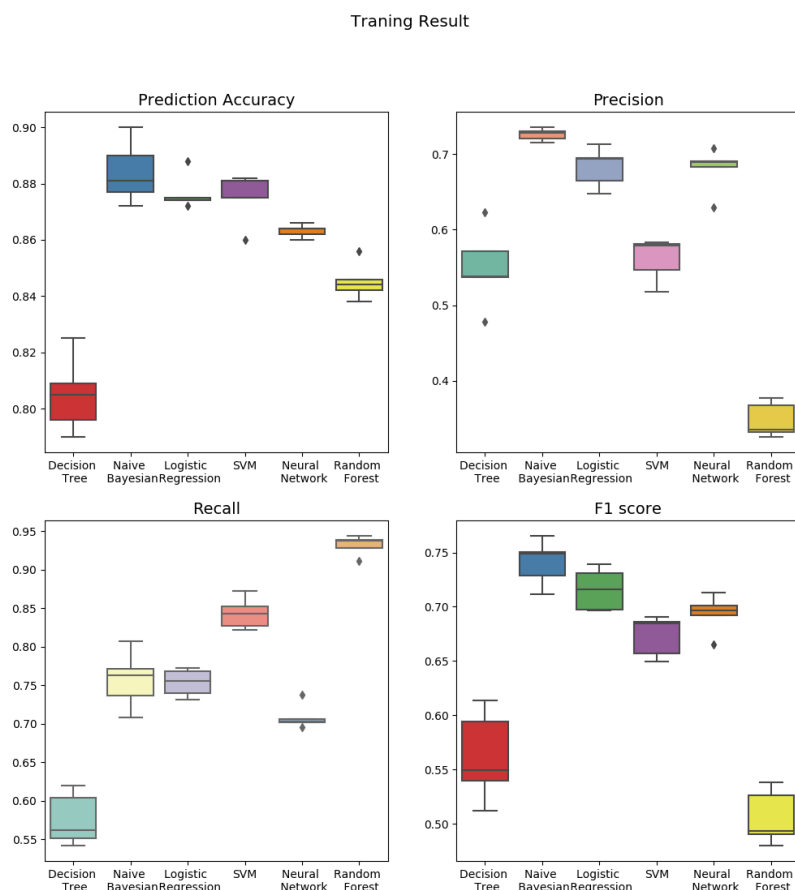


图 3-17 模型训练比较结果(code_ML_result.png)

3.5.2 方法二：使用命令提示符运行系统

和 3.5.1 节过程一致,用任意文本文件打开代码,将 PREDICTION_ONLY 由 True 修改为 False, 注意第一个字母大写, 如图 3-18 所示。

在训练中, 命令提示符对话框内会输出训练时间 (53.246s), 训练结果 (winning model is: NB objective(F1 score)= 0.7409658610738933)

同样的, code_ML_result.png (图 3-17) 和训练模型 (图 3-16) 会被输出。

```
# Author: Xiaoyu Li
# Last Updated: 2019-12-16

...

Input all configuration parameters (modified by user)
...

PREDICTION_ONLY = False
predict_file_name = "predict_text.xlsx"
train_file_name = "train_text.xlsx"
dictionary_file_name = "dictionary.xlsx"
model_name = "code_model_MultinomialNB.sav"
```

图 3-18 用 txt 文本文件打开 Python 代码

```
====finished loop 4 it takes time 53.246 s====
model= tree Mean of F1_result0 is: 0.5618510701480389 key_winner= svm
model= NB Mean of F1_result0 is: 0.7409658610738933 key_winner= tree
model= LR Mean of F1_result0 is: 0.7161463489412864 key_winner= NB
model= svm Mean of F1_result0 is: 0.6736304250753347 key_winner= NB
model= nn Mean of F1_result0 is: 0.6937181823806429 key_winner= NB
model= RF Mean of F1_result0 is: 0.5057041922423926 key_winner= NB
*****winning model is: NB objective(F1 score)= 0.7409658610738933 *****
```

图 3-19 运行本软件进行模型训练(通过命令提示符)