

## Instalação NLTK Windows 64-bit

1. Instalar o Python mais atual em <https://www.python.org/downloads/>;
  - a. Após a instalação do Python, colocar a pasta do Python (Python34) e a pasta Scripts (Python34\Scripts) no path do sistema.
  - b. O instalador do Python tem uma opção que faz isso automaticamente. Na tela de customização, role a barra de rolagem até o final, clique em *Add python.exe to Path* e selecione a opção *Entire feature will be installed on local hard drive*. (Figura 1)

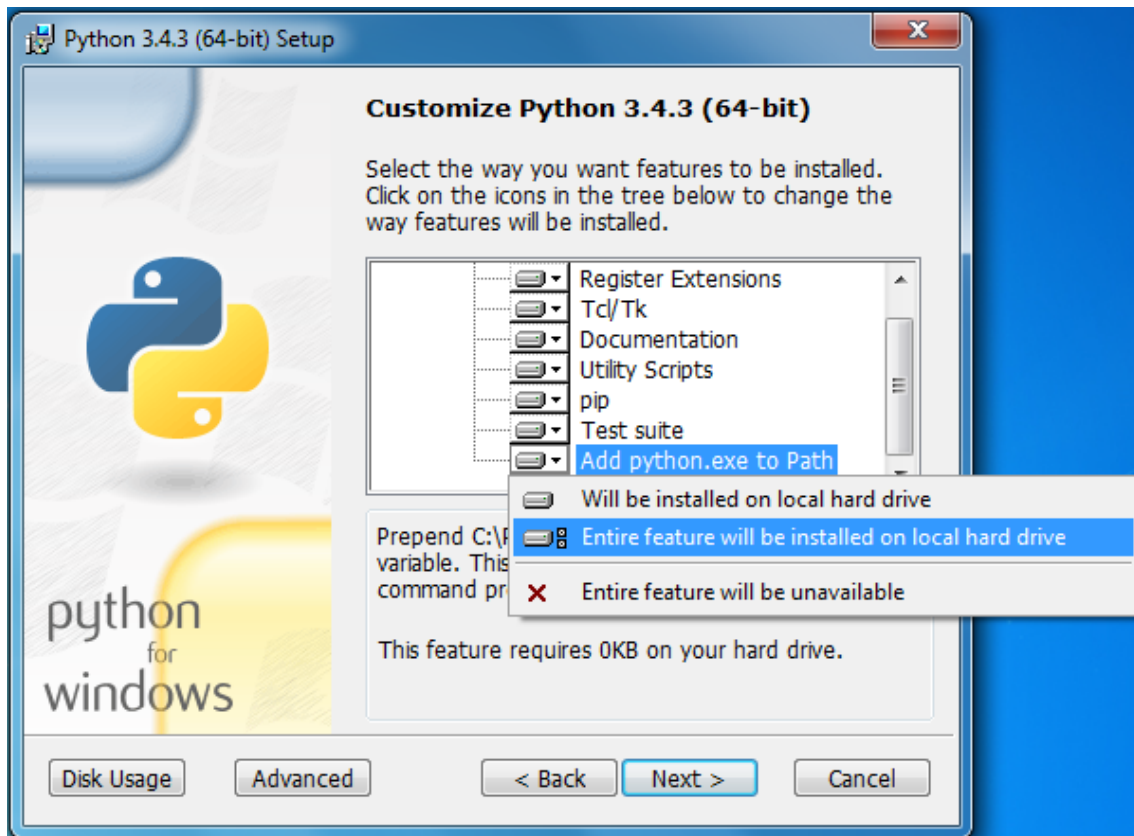


Figura 1: Adicionando o Python no Path do sistema pelo instalador.

2. Baixar os arquivos do NTLK, Numpy e PyWin32 através dos links:
  - a. NTLK - <http://www.lfd.uci.edu/~gohlke/pythonlibs/#nltk>
  - b. Numpy - <http://www.lfd.uci.edu/~gohlke/pythonlibs/#numpy>
  - c. PyWin32 - <http://www.lfd.uci.edu/~gohlke/pythonlibs/#pywin32>
  - d. PyYAML - <http://www.lfd.uci.edu/~gohlke/pythonlibs/#pyyaml>

**Obs 1:** Recomendável salvar os arquivos dentro de uma pasta chamada NLTK.

**Obs 2:** No Windows 10 baixar os arquivos referentes à arquitetura de 32-bit.

3. Abrir o prompt de comando, acessar a pasta onde estão salvos os arquivos baixados no tópico 2 instalar os arquivos através do comando

```
pip3.4 install <nome_do_arquivo>
```

Exemplo:

```
pip3.4 install nltk-3.0.1-py2.py3-none-any.whl
```

4. Abrir o IDLE Python GUI:

- a. Digitar `import nltk`, <ENTER>
  - b. Digitar `nltk.download()`, <ENTER>
5. Na tela de download do NLTK, indique em *Download Directory* a pasta **nltk\_data** (Figura 2), selecione a linha com *Identifier all* e clique em *Download*. Os pacotes serão baixados. A Figura 3 mostra a tela do NLTK Downloader quando os pacotes forem todos baixados.

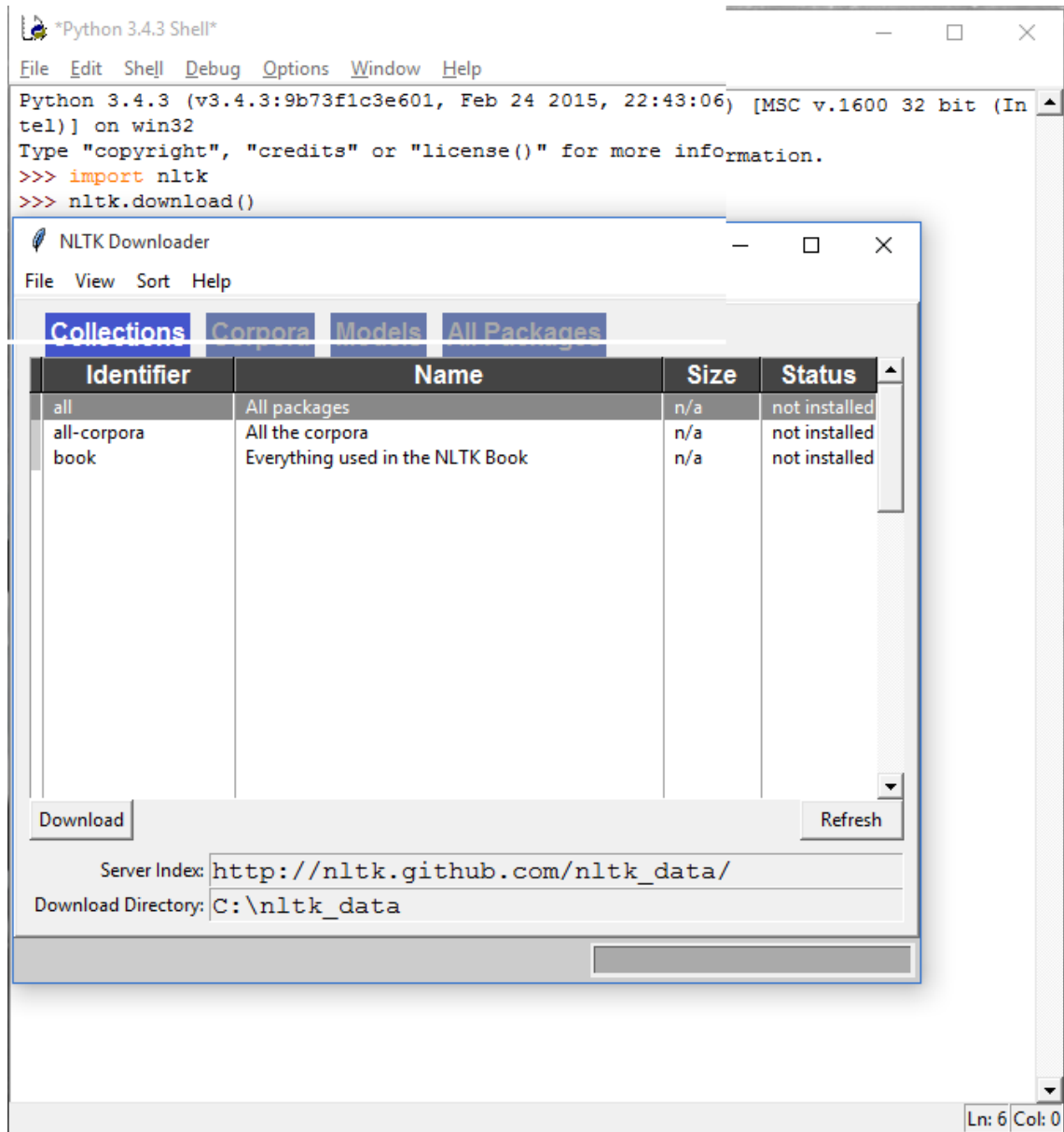


Figura 2: NLTK downloader.

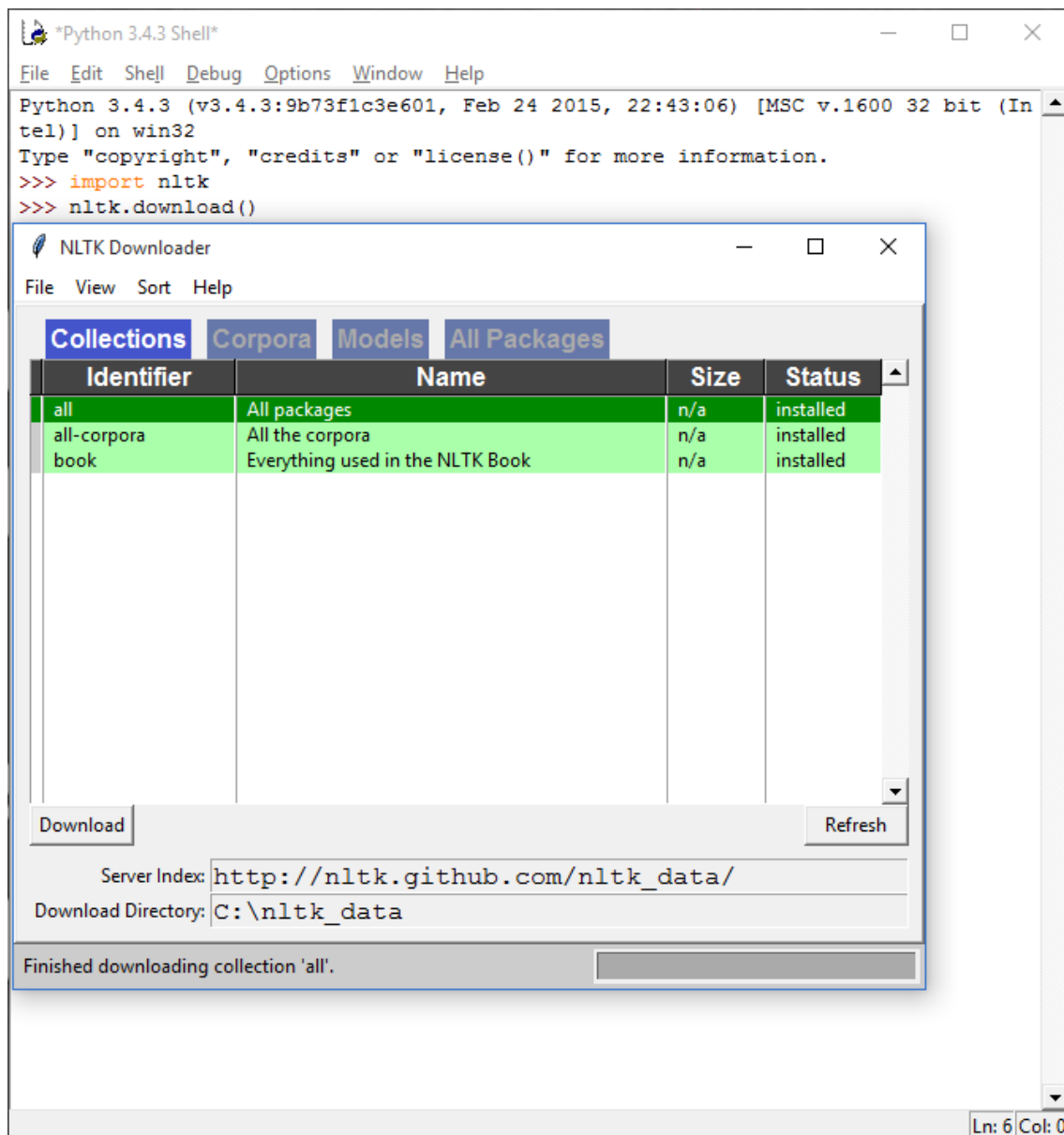
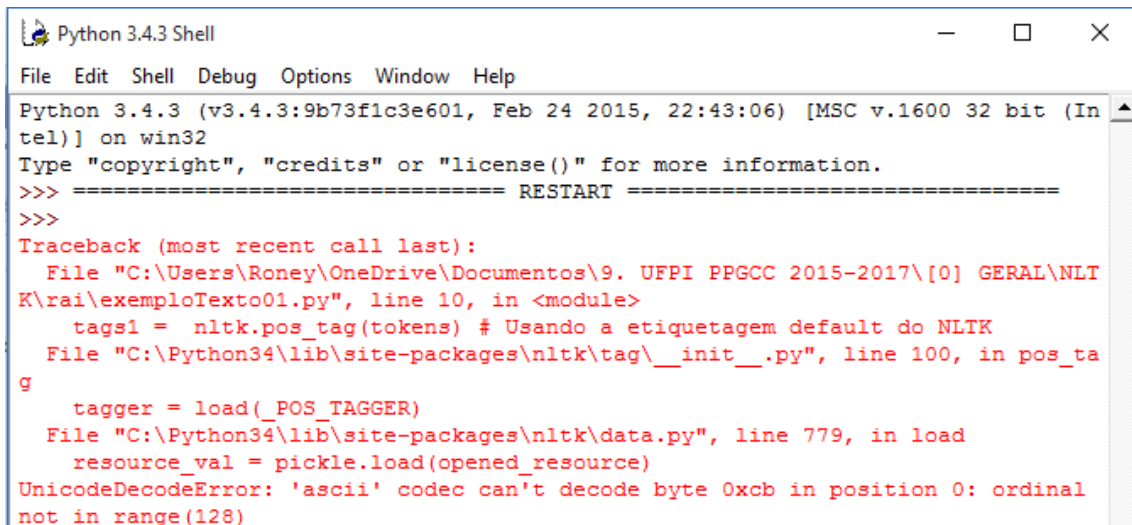


Figura 3: NLTK Downloader quando os pacotes forem baixados.

6. Baixe [aqui](#) os arquivos de teste no NLTK feitos pelo grupo. Eles serão utilizados para terminar a configuração do NLTK Python.
7. Abra novamente o IDLE Python GUI, vá em File > Open e abra o arquivo **exemploBigramas.py**. Pressione F5 para executar o código. Um erro como o da Figura 4 vai aparecer no *shell*. Para corrigir:
  - a. Abra o arquivo data.py que se encontra em **Python34\Lib\site-packages\nltk**;
  - b. Navegue pelo arquivo até a linha 779. Altere o que existe lá por **resource\_val = pickle.load(opened\_resource, encoding='iso-8859-1')**. [Fonte: [Stackoverflow](#)]
  - c. Execute novamente o arquivo de teste pressionando F5. Outro erro será mostrado. Esse erro é referente ao módulo pylab que mostra gráficos de diversos tipos e as *parse trees*. Para corrigi-lo, basta fazer o download do Matplotlib em <http://www.lfd.uci.edu/~gohlke/pythonlibs/#matplotlib>,

salvá-lo na pasta criada no tópico 2 e instalar seguindo os mesmos requisitos do tópico 3.

Obs: No Windows 10 baixar os arquivos referentes à arquitetura de 32-bit.



```
Python 3.4.3 Shell
File Edit Shell Debug Options Window Help
Python 3.4.3 (v3.4.3:9b73f1c3e601, Feb 24 2015, 22:43:06) [MSC v.1600 32 bit (Intel)] on win32
Type "copyright", "credits" or "license()" for more information.
>>> ===== RESTART =====
>>>
Traceback (most recent call last):
  File "C:\Users\Roney\OneDrive\Documentos\9. UFPI PPGCC 2015-2017\[0] GERAL\NLTK\rai\exemploTexto01.py", line 10, in <module>
    tags1 = nltk.pos_tag(tokens) # Usando a etiquetagem default do NLTK
  File "C:\Python34\lib\site-packages\nltk\tag\__init__.py", line 100, in pos_tag
    tagger = load(_POS_TAGGER)
  File "C:\Python34\lib\site-packages\nltk\data.py", line 779, in load
    resource_val = pickle.load(opened_resource)
UnicodeDecodeError: 'ascii' codec can't decode byte 0xcb in position 0: ordinal not in range(128)
```

Figura 4: Erro na execução do primeiro programa do NLTK no Python.

- d. Execute mais uma vez o arquivo teste.
8. Pronto! O NLTK Python está instalado e configurado com as premissas iniciais para começar a trabalhar. Outras ferramentas podem ser adicionadas no decorrer do aprendizado e da necessidade que se tenha.