**[Nicholas Drabowski](mailto:nicholas@portaltelemedicina.com.br)**[<nicholas@portaltelemedicina.com.br>](mailto:nicholas@portaltelemedicina.com.br)

Para

[alfredorp33@yahoo.com.br](mailto:alfredorp33@yahoo.com.br)

 Mai 7 em 4:50 PM

[Boxbe](https://www.boxbe.com/overview)  Nicholas Drabowski (nicholas@portaltelemedicina.com.br) is not on [your Guest List](https://www.boxbe.com/approved-list?tc_serial=49613751996&tc_rand=1785248247&utm_source=stf&utm_medium=email&utm_campaign=ANNO_MWTP&utm_content=001&key=1CeKYfX8yzHgtiINJ%2BnBJQNTd73dxJbL3a1Bmr0jzBY%3D&token=2S43rBg97dZnwTyFQhqw3BC0gwgYyjPp6gZatzppQDd8cK4Z%2Frlqy1QXwxCxk1Q0) | [Approve sender](https://www.boxbe.com/anno?tc_serial=49613751996&tc_rand=1785248247&utm_source=stf&utm_medium=email&utm_campaign=ANNO_MWTP&utm_content=001&key=1CeKYfX8yzHgtiINJ%2BnBJQNTd73dxJbL3a1Bmr0jzBY%3D&token=2S43rBg97dZnwTyFQhqw3BC0gwgYyjPp6gZatzppQDd8cK4Z%2Frlqy1QXwxCxk1Q0) | [Approve domain](https://www.boxbe.com/anno?tc_serial=49613751996&tc_rand=1785248247&utm_source=stf&utm_medium=email&utm_campaign=ANNO_MWTP&utm_content=001&dom&key=1CeKYfX8yzHgtiINJ%2BnBJQNTd73dxJbL3a1Bmr0jzBY%3D&token=2S43rBg97dZnwTyFQhqw3BC0gwgYyjPp6gZatzppQDd8cK4Z%2Frlqy1QXwxCxk1Q0)

Boa tarde Alfredo,

Segue o teste para cientista de dados, e o arquivo necessário para o teste em anexo. Você tem até o dia 16/05/2019 23:59 para entregar o teste.

Não compartilhe com ninguém que não tenha o e-mail @[portaltelemedicina.com.br](http://portaltelemedicina.com.br/)

Não esqueça do read-me descrevendo perfeitamente como configurar o projeto.

Boa sorte!

**Test for Data scientist fapesp SIAGeDI - Portal Telemedicina**

The 'test\_data\_CANDIDATE.csv' file contains some patient's data. The test consists in analyzing the data, building a sex predictor and a script that allows the use of the model. The deliverables of the test are:

1. The analysis of the data

- can be presented in any form, but a python notebook is preferable

- an analysis and/or brief explanation of the model delivered in the item 2 must also be present.

2. A model that predicts whether a batch of patients are male or female, given an entirely new CSV file containing the same columns, except the 'sex' column.

- the script must be a python file named 'sex\_predictor.py'

- a 'requirements.txt' file is expected, so that the necessary dependencies can be installed in a virtual environment using the command 'pip install requirements.txt'

- the script will be used this way:

(env) python sex\_predictor.py --input\_file newsample.csv

- the newsample.csv file is a file with the same structure of the 'test\_data\_CANDIDATE.csv', but with an unknown number of lines (n)

- the expected output is a table saved in the same directory but with the name 'newsample\_PREDICTIONS\_{candidate\_fullname}.csv' containing only the column 'sex', which has the predicted values.

A compressed folder containing the files can be delivered by e-mail.

**Columns descriptions:**

age: in years

sex: (M = male; F = female)

cp: chest pain type

trestbps: resting blood pressure (in mm Hg on admission to the hospital)

chol: serum cholesterol in mg/dl

fbs: (fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)

restecg: resting electrocardiographic results

thalach: maximum heart rate achieved

nar: number of arms

exang: exercise induced angina (1 = yes; 0 = no)

oldpeak: ST depression induced by exercise relative to rest

slope: the slope of the peak exercise ST segment

hc: patient's hair colour

sk: patient's skin colour

trf: time spent in traffic daily (in seconds)

ca: number of major vessels (0-3) colored by flourosopy

thal: 3 = normal; 6 = fixed defect; 7 = reversable defect

--

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| https://ecp.yusercontent.com/mail?url=http%3A%2F%2Fportaltelemedicina.com.br%2Fwp-content%2Fuploads%2F2017%2F12%2Fnicholas-21.png&t=1557446555&ymreqid=4a53fb40-3794-0636-0121-120402010000&sig=4eDjSoGyYs1IODeXAbFyOQ--~C |  | |  |  | | --- | --- | | https://ecp.yusercontent.com/mail?url=http%3A%2F%2Fportaltelemedicina.com.br%2Fwp-content%2Fuploads%2F2017%2F11%2Fassinatura-06.png&t=1557446555&ymreqid=4a53fb40-3794-0636-0121-120402010000&sig=OqtJO2sMoZoVJVmxmahaWg--~C | **NICHOLAS DRABOWSKI** | |  | *R&D ENGINEER | DATA SCIENTIST* | | https://ecp.yusercontent.com/mail?url=http%3A%2F%2Fportaltelemedicina.com.br%2Fwp-content%2Fuploads%2F2017%2F11%2Fassinatura-07.png&t=1557446555&ymreqid=4a53fb40-3794-0636-0121-120402010000&sig=lmMv7l.2Ag8jRw9f0E.8qA--~C | +55 (48) 99682-8912 | |  | +55 (48) 3331-3502 | | [https://ecp.yusercontent.com/mail?url=http%3A%2F%2Fportaltelemedicina.com.br%2Fwp-content%2Fuploads%2F2017%2F11%2Fassinatura-08.png&t=1557446555&ymreqid=4a53fb40-3794-0636-0121-120402010000&sig=k31LbjlM4OHLONJ2IhNlWA--~C](https://www.linkedin.com/in/ndrabowski) | /ndrabowski | | https://ecp.yusercontent.com/mail?url=http%3A%2F%2Fportaltelemedicina.com.br%2Fwp-content%2Fuploads%2F2017%2F11%2Fassinatura-09.png&t=1557446555&ymreqid=4a53fb40-3794-0636-0121-120402010000&sig=rzhMu1N_uRin7fpYSJulPA--~C | qg\_nicholas | |
| https://ecp.yusercontent.com/mail?url=http%3A%2F%2Fportaltelemedicina.com.br%2Fwp-content%2Fuploads%2F2017%2F12%2Fassinatura-01-1.png&t=1557446555&ymreqid=4a53fb40-3794-0636-0121-120402010000&sig=oOHk1avz3UC5ynKULimffw--~C |  | [https://ecp.yusercontent.com/mail?url=http%3A%2F%2Fportaltelemedicina.com.br%2Fwp-content%2Fuploads%2F2017%2F12%2Fassinatura-02-1.png&t=1557446555&ymreqid=4a53fb40-3794-0636-0121-120402010000&sig=VXy7MPilIb4k1UoMjPfS1g--~C](http://portaltelemedicina.com.br/)[https://ecp.yusercontent.com/mail?url=http%3A%2F%2Fportaltelemedicina.com.br%2Fwp-content%2Fuploads%2F2017%2F12%2Fassinatura-03-1.png&t=1557446555&ymreqid=4a53fb40-3794-0636-0121-120402010000&sig=bif4gR3_xnmSlCZQaBJFzg--~C](https://www.youtube.com/channel/UC0KH36-OXHFIKjlRY2GyAtQ)[https://ecp.yusercontent.com/mail?url=http%3A%2F%2Fportaltelemedicina.com.br%2Fwp-content%2Fuploads%2F2017%2F12%2Fassinatura-04-1.png&t=1557446555&ymreqid=4a53fb40-3794-0636-0121-120402010000&sig=p5quAYxbLVoQOp0vgrxRGw--~C](https://pt-br.facebook.com/PortalTelemedicina/)[https://ecp.yusercontent.com/mail?url=http%3A%2F%2Fportaltelemedicina.com.br%2Fwp-content%2Fuploads%2F2017%2F12%2Fassinatura-15.png&t=1557446555&ymreqid=4a53fb40-3794-0636-0121-120402010000&sig=QrzFzS1WG5F7S7hnZh598Q--~C](https://www.linkedin.com/company/9426084/) |

* test\_data\_CANDIDATE(2)