Using Andersen’s model of health care utilization to assess factors associated with COVID-19 testing among adults in Bangladesh: an online survey

Ehsan Suez

2021-11-12

# 1 Summary/Abstract

I’ll write it later.,

# 2 Introduction

Testing can help people determine if they are infected with SARS-CoV-2 – regardless of whether they have symptoms – and whether they are at risk of spreading the infection to others. Testing of all people for SARS-CoV-2, including those who have no symptoms, who show symptoms of infection, and who may have been exposed to the virus can help prevent the spread of COVID-19 and identify people in need of care in a timely fashion. A positive test early in the course of the illness enables individuals to isolate themselves – reducing the chances of infecting others and allowing them to seek treatment earlier, likely reducing disease severity and the risk of long-term disability, or death.

A negative test does not mean the person will not be able to acquire and transmit the infection in the future. Therefore, even if an individual is test negative, continuation of hygiene practices to self-protect and others by washing hands with soaps/ sanitizers frequently, physical distancing, and wearing a face mask is required. A positive test means isolation is mandatory, and that others with whom the individual may have been in contact since the time of exposure should also get tested. Since it is recognized that nearly half of all SARS-CoV-2 infections are transmitted by pre-symptomatic and asymptomatic people, identifying infected individuals while they are pre-symptomatic, as well as those who are asymptomatic, will play a major role in containing the pandemic.

Scanty research has been conducted to identify factors associated with the uptake of COVID-19 testing. In a cross-sectional community survey in Ontario, Canada, 3.3% tested positive for SARS-CoV-2. The odds of being tested increased with age, male sex, several underlying health conditions, previous use of health care services, and higher household income (Sundaram et al., 2021). Comparing the odds of COVID-19 positive diagnosis with test-negatives and untested individuals, older age, certain comorbidities, such as hypertension, diabetes and heart diseases, higher previous use of health care services were associated with increased odds of a positive SARS-CoV-2 test result (Sundaram et al., 2021). Using UK biobank data, being tested for COVID-19 was associated with male sex, Black ethnicity, lower socioeconomic status, occupation (being a health care worker, retired, unemployed), ever smoker, comorbidities, and exposure to particulate matter (PM) 2.5 absorbance (Chadeau-Hyam et al., 2020). Among tested individuals, only male sex, non-White ethnicity, and lower education and none of the comorbidities or health risk factors were associated with a COVID-19 positive diagnosis [Chadeau-Hyam]. Workers in the healthcare sector are also at a higher risk of testing positive. A meta-analytic study showed that the pooled prevalence of healthcare workers testing positive for COVID-19 was between 7 to 11% (Gómez-Ochoa et al., 2020).

## 2.1 General Background Information

This study aimed to investigate, using the Andersen’s model of health care utilization, factors associated with COVID-19 testing among adults in Bangladesh.

## 2.2 Description of data and data source

This will be a descriptive cross-sectional online study. To get data for this study, between 10 December 2020 and 9 February 2021 an online survey was organized in Bangladesh. In total 759 adults (median age 45 years, interquartile range 33-57 years, range 18-93 years), participated in the study. COVID-19 testing/infection status was assessed by self-report.

## 2.3 Questions/Hypotheses to be addressed

The main question is to find out the key factors associated with covid-19 testing

# 3 Methods and Results

***Study design, sample, and procedure***

This will be a descriptive cross-sectional online study conducted in Bangladesh between 10 December 2020 to 9 February 2021. Participant inclusion criteria were 18 years and older, any gender, and provision of electronic informed consent.

***Measures*** Using Andersen’s model of health care utilization , study variables will be categorized into outcome variable, predisposing factors, enabling/disabling factors and need for care factors.

***Outcome variable*** COVID-19 testing/infection status will be assessed with the question, “Since the beginning of the COVID-19 outbreak, do you have information on your infection status?” Response options were 1=not tested/does not know test results, 2=negative, and 3=positive.

***Predisposing factors*** Sociodemographic factors included age, sex, country of residence, educational level, and the (estimated) age(s) of their housemate(s). Chronic/underlying diseases includs heart disease, hypertension, diabetes, cancer, HIV, tuberculosis, and chronic asthma; coded as “0” none and “1” at least presence of one clinically diagnosed condition.

***COVID-19 preventive measures*** Participants were asked, “During the past 7 days, have you been observing any of the following preventive measures against COVID-19? 1) Social distancing of at least 1.5m, 2) Wearing a face mask, 3) Hand hygiene (regular handwashing with soap or using hand gel), and 4) Coughing hygiene (covering the mouth when coughing or sneezing). A composite non-adherence to all four COVID-19 preventive measure was calculated by coding each negative response with “1,” summing scores ranging from 0-4 (Cronbach’s alpha 0.7).

***Enabling factors*** Enabling factors include self-perceived socio-economic status, self-perceived area of residence, being a student or worker in the health care sector, source of COVID-19 information/advice most trusted (coded as other, including family and friends, radio/TV, social media, religious authorities and health personnel(Bono et al., 2021).

***Need for care factors*** include two questions on 1) the level of fear/worry of being infected with COVID-19 (ranging from 1=not at all worried to 5=extremely worried), and 2) having been quarantined (either at home or elsewhere) at any point in time during the COVID-19 epidemic.

***Data analysis***

Descriptive statistics will be used to describe the study population. Logistic regression will be used to assess associations between predisposing factors, enabling and disabling factors, need of care factors and COVID-19 testing status, COVID-19 positive versus negative status and COVID-19 positive versus negative and not tested status. Variables significant at <0.05 in univariate analyses were subsequently included in the multivariable logistic regression models. Statistical analyses will conducted using R.

## 3.1 Data aquisition

*As applicable, explain where and how you got the data. If you directly import the data from an online source, you can combine this section with the next.*

## 3.2 ## Data import and cleaning\_ html\_document

library(here)

## here() starts at /Users/ehsansuez/Study /MADA\_2021/EhsanSuez-MADA-project

## 3.3 Exploratory analysis

Table 3.1: **Overall Descriptive Statistics of the Participants**

|  |  |
| --- | --- |
| X. | Total |
|  | (N=706) |
| Age |  |
| Mean (SD) | 30.7 (10.8) |
| Median [Min, Max] | 27.0 [18.0, 110] |
| Sex |  |
| Male | 387 (54.8%) |
| Female | 319 (45.2%) |
| education |  |
| University Postgraduate degree holder | 168 (23.8%) |
| University Undergraduate degree holder | 383 (54.2%) |
| Secondary | 149 (21.1%) |
| Primary | 6 (0.8%) |
| economic\_status |  |
| Low income category | 54 (7.6%) |
| Upper middle income category | 348 (49.3%) |
| Lower middle income category | 259 (36.7%) |
| High income category | 45 (6.4%) |
| student\_or\_heathworker |  |
| Yes | 226 (32.0%) |
| No | 480 (68.0%) |
| residence |  |
| An urban setting / city / town | 445 (63.0%) |
| A Sub-urban setting / urban slum | 197 (27.9%) |
| A rural place / village | 64 (9.1%) |
| trusted\_source |  |
| Other | 32 (4.5%) |
| Health personnel | 237 (33.6%) |
| Social Media (WhatsApp, Facebook, Twitter, etc) | 151 (21.4%) |
| Radio / TV | 217 (30.7%) |
| Family and friends | 21 (3.0%) |
| None of the above | 32 (4.5%) |
| Religious authorities (Pastor, Priest, Imam, etc) | 16 (2.3%) |
| work\_from\_home |  |
| Yes | 359 (50.8%) |
| No | 347 (49.2%) |
| quarantined\_or\_not |  |
| Yes | 407 (57.6%) |
| No | 299 (42.4%) |
| infection\_status\_info |  |
| noTest | 445 (63.0%) |
| negative | 197 (27.9%) |
| positive | 64 (9.1%) |
| worry\_level |  |
| 3=Moderately worried | 230 (32.6%) |
| 2=A little worried | 256 (36.3%) |
| 4=Very worried | 77 (10.9%) |
| 5=Extremely worried | 18 (2.5%) |
| 1=Not at all worried | 125 (17.7%) |
| chronic\_disease\_history |  |
| I dont have #any\_of\_the\_previously\_mentioned Disease | 554 (78.5%) |
| X- I have #at\_least\_one\_previously\_mentioned Disease | 152 (21.5%) |
| social\_distance\_practice |  |
| No, I don’t followed this measure | 458 (64.9%) |
| X- Yes I followed this measure | 248 (35.1%) |
| follow\_preventive\_measure |  |
| X-Yes, I followed #any\_of\_the\_previously\_mentioned measures | 689 (97.6%) |
| No, I don’t followed #any\_of\_the\_previously\_mentioned measures | 17 (2.4%) |

As the infection status information of the participants is the main outcome variable, let’s see its distribution among the participants:

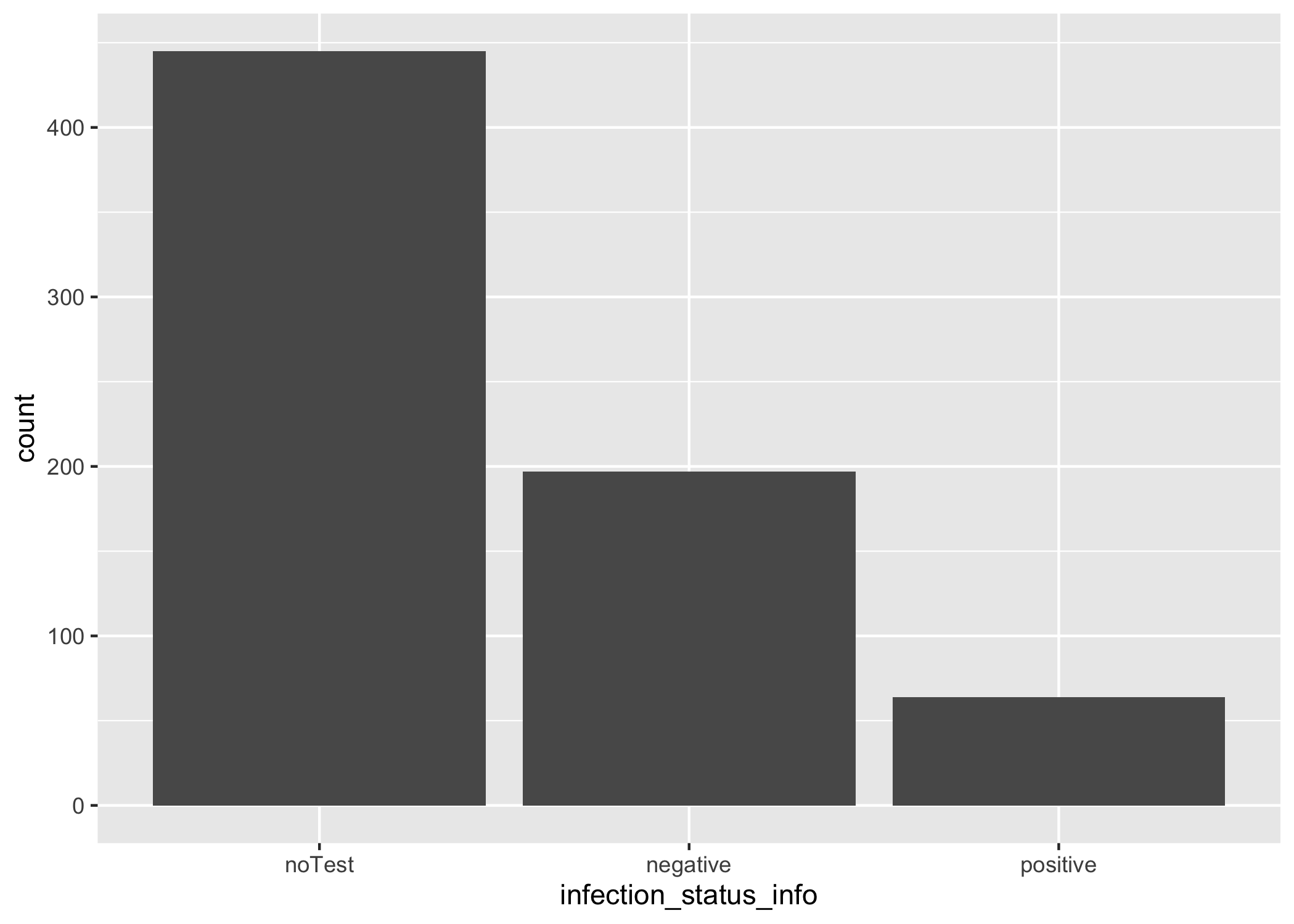


Figure 3.1: Infection Status Information.

Among the participants, **78.9%** are less than or equal to 35 years old. That indicates that majority of the participants tend to be younger. While **45.18%** of the participants were female, **54.82%** were male, which shows less chance of biad due to sex. **78.05%** of the participants had a bachelor or post-bachelor degree, clearly indicating that majority of the participants had pretty good education background. Almost **50%** of the participants belongs to “upper middle income category,” indicating that at least half of the participants don’t belong to the poorer side. Less than **10%** of the participants live in a village/rural area. Health professionals and Tv/radio seems to be the most trusted source to the participants. Apparently equal half of the participants worked from home or not at some stage of the study. Major half of the participants did quarantine at some stage during the study. **63%** of the participants didn’t go through a COVID test. Almost half of the participants **(46%)** were moderate or more worried about reinfection. Only **22%** of the participants recorded to have some sort of chronic disease. **(Separate analysis for each variable was done and the data can be accessed/viewd through the “supplementary.rmd” file)**

Let’s look at the logistic regression analysis and its statistics summary between the **outcome variable** (infection status info) and the three key **predisposing factors** (Age, Sex, Education level).

Table 3.2: logistic regression betweem infection status and Key Predisposing Factors

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| (Intercept) | -1.7832470 | 0.3331593 | -5.3525361 | 0.0000001 |
| Age | 0.0458298 | 0.0080757 | 5.6750136 | 0.0000000 |
| SexFemale | 0.4033255 | 0.1635476 | 2.4661042 | 0.0136592 |
| educationUniversity Undergraduate degree holder | -0.4111288 | 0.1970401 | -2.0865232 | 0.0369313 |
| educationSecondary | -0.6900496 | 0.2455377 | -2.8103608 | 0.0049486 |
| educationPrimary | -0.4371349 | 0.9496023 | -0.4603347 | 0.6452760 |

Table 3.3: logistic regression statistics betweem infection status and Key Predisposing Factors

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| null.deviance | df.null | logLik | AIC | BIC | deviance | df.residual | nobs |
| 930.211 | 705 | -436.7208 | 885.4416 | 912.7993 | 873.4416 | 700 | 706 |

Next, we want to see the logistic regression analysis statistics between the **outcome variable** (infection status info) and the three key **enabling factors** (socio economic status, area of residence,student/healthworker, source of covid news).

Table 3.4: logistic regression betweem infection status and Key Enabling Factors

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| (Intercept) | -1.3559366 | 0.5345008 | -2.5368281 | 0.0111862 |
| economic\_statusUpper middle income category | 1.4851753 | 0.4287608 | 3.4638785 | 0.0005324 |
| economic\_statusLower middle income category | 0.9112994 | 0.4350794 | 2.0945589 | 0.0362102 |
| economic\_statusHigh income category | 1.9065113 | 0.5148440 | 3.7030857 | 0.0002130 |
| residenceA Sub-urban setting / urban slum | 0.1044858 | 0.1829930 | 0.5709824 | 0.5680116 |
| residenceA rural place / village | -0.8666444 | 0.3695104 | -2.3453858 | 0.0190074 |
| trusted\_sourceHealth personnel | -0.1196328 | 0.3989853 | -0.2998426 | 0.7642972 |
| trusted\_sourceSocial Media (WhatsApp, Facebook, Twitter, etc) | -0.5401488 | 0.4172613 | -1.2945096 | 0.1954894 |
| trusted\_sourceRadio / TV | -0.5963396 | 0.4061255 | -1.4683628 | 0.1420057 |
| trusted\_sourceFamily and friends | -0.3366449 | 0.6110896 | -0.5508929 | 0.5817071 |
| trusted\_sourceNone of the above | -0.3203110 | 0.5289505 | -0.6055596 | 0.5448072 |
| trusted\_sourceReligious authorities (Pastor, Priest, Imam, etc) | -0.5074121 | 0.6722440 | -0.7548035 | 0.4503669 |

Table 3.5: logistic regression statistics betweem infection status and Key enabling Factors

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| null.deviance | df.null | logLik | AIC | BIC | deviance | df.residual | nobs |
| 930.211 | 705 | -440.6209 | 905.2419 | 959.9573 | 881.2419 | 694 | 706 |

Next, we want to see the logistic regression analysis statistics between the **outcome variable** (infection status info) and the level of fear of infection and quarantined or not **(Key need for care factors )**

Table 3.6: logistic regression betweem infection status and Key Need for care Factor

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| term | estimate | std.error | statistic | p.value |
| (Intercept) | -0.0196069 | 0.1484504 | -0.1320768 | 0.8949235 |
| quarantined\_or\_notNo | -1.0814712 | 0.1754811 | -6.1628928 | 0.0000000 |
| worry\_level2=A little worried | -0.1796001 | 0.1927321 | -0.9318638 | 0.3514069 |
| worry\_level4=Very worried | 0.6275857 | 0.2746511 | 2.2850287 | 0.0223111 |
| worry\_level5=Extremely worried | 0.6861161 | 0.5101150 | 1.3450225 | 0.1786180 |
| worry\_level1=Not at all worried | -0.8368647 | 0.2718371 | -3.0785518 | 0.0020801 |

Table 3.7: logistic regression statistics betweem infection status and Key Need for Care Factors

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| null.deviance | df.null | logLik | AIC | BIC | deviance | df.residual | nobs |
| 930.211 | 705 | -428.6532 | 869.3064 | 896.6641 | 857.3064 | 700 | 706 |

The comparison between these three models give us this information:

Table 3.8: Comparison of the Models

|  |  |  |  |
| --- | --- | --- | --- |
| Resid. Df | Resid. Dev | Df | Deviance |
| 700 | 873.4416 | NA | NA |
| 694 | 881.2419 | 6 | -7.800298 |
| 700 | 857.3064 | -6 | 23.935522 |

we want to predict each infection status level, not a single contrast by using multinom\_reg(). The result is as following:

Table 3.9: Metrices of Training Data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| .metric | .estimator | mean | n | std\_err | .config |
| accuracy | multiclass | 0.6207446 | 25 | 0.0069295 | Preprocessor1\_Model1 |
| roc\_auc | hand\_till | 0.6736977 | 25 | 0.0084538 | Preprocessor1\_Model1 |

Table 3.10: Metrices of Testing Data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| .metric | .estimator | mean | n | std\_err | .config |
| accuracy | multiclass | 0.5744031 | 25 | 0.0134971 | Preprocessor1\_Model1 |
| roc\_auc | hand\_till | 0.5972284 | 25 | 0.0186692 | Preprocessor1\_Model1 |

# 4 Discussion

## 4.1 Summary and Interpretation

*Summarize what you did, what you found and what it means.*

## 4.2 Strengths and Limitations

*Discuss what you perceive as strengths and limitations of your analysis.*

## 4.3 Conclusions

*What are the main take-home messages?*

# References

Bono, S. A., Faria de Moura Villela, E., Siau, C. S., Chen, W. S., Pengpid, S., Hasan, M. T., … Colebunders, R. (2021). Factors affecting COVID-19 vaccine acceptance: An international survey among low- and middle-income countries. *Vaccines*, *9*(5). <https://doi.org/10.3390/vaccines9050515>

Chadeau-Hyam, M., Bodinier, B., Elliott, J., Whitaker, M. D., Tzoulaki, I., Vermeulen, R., … Elliott, P. (2020). Risk factors for positive and negative COVID-19 tests: a cautious and in-depth analysis of UK biobank data. *International Journal of Epidemiology*, *49*(5), 1454–1467. <https://doi.org/10.1093/ije/dyaa134>

Gómez-Ochoa, S. A., Franco, O. H., Rojas, L. Z., Raguindin, P. F., Roa-Díaz, Z. M., Wyssmann, B. M., … Muka, T. (2020). COVID-19 in Health-Care Workers: A Living Systematic Review and Meta-Analysis of Prevalence, Risk Factors, Clinical Characteristics, and Outcomes. *American Journal of Epidemiology*, *190*(1), 161–175. <https://doi.org/10.1093/aje/kwaa191>

Sundaram, M. E., Calzavara, A., Mishra, S., Kustra, R., Chan, A. K., Hamilton, M. A., … Kwong, J. C. (2021). Individual and social determinants of SARS-CoV-2 testing and positivity in ontario, canada: A population-wide study. *CMAJ*, *193*(20), E723–E734. <https://doi.org/10.1503/cmaj.202608>