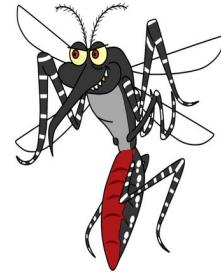


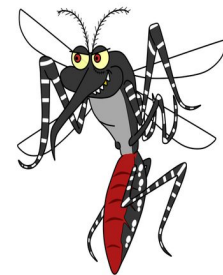
Casos de Dengue na cidade do Recife em 2016

IF706 - Ciência dos Dados



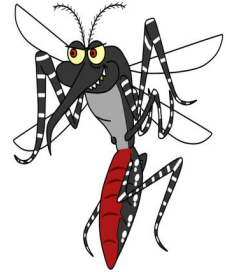
Equipe

- Antônio Portela (acpr)
- Pedro Moraes (phsm)
- Thiago Aquino (tas4)



Roteiro

- Introdução
- Limpeza dos dados
- Pré Processamento dos dados
- Análises
- Resultado por categoria
- Sintoma X Categoria
- Curva ROC
- Matriz de Confusão
- Avaliação dos Classificadores



Introdução

A dengue é uma das arboviroses que mais afligem a população brasileira. Por ser transmitida através de mosquitos, a população, que sofre pela falta de saneamento básico, é muito afetada pela doença. Este trabalho estuda os casos de dengue registrados em Recife no ano de 2016. Objetivamos estudar se locais com alta incidência de casos de dengue estão relacionados a uma quantidade maior ou menor de unidades de saúde de forma a verificar se elas exercem alguma influência nos casos. Além disso, estudaremos as relações entre os sintomas do caso com a gravidade da doença, para, a partir disso, verificar se existe um conjunto de sintomas específicos que está fortemente relacionado a gravidade da doença.



Limpeza dos dados

	nu_notificacao	tp_notificacao	co_cid	dt_notificacao	ds_semana_notificacao	notificacao_ano	co_uf_notificacao	co_municipio_notificacao	id_regional	c
7277	2591534	2	A90	14-03-2016	201611	2016	26	261160	1497.0	
576	2586902	2	A90	05-10-2016	201619	2016	26	261160	1497.0	
2272	2291568	2	A90	01-04-2016	201601	2016	26	261160	1497.0	
5336	2157221	2	A90	18-02-2016	201607	2016	26	261160	1497.0	
4641	2586297	2	A90	15-06-2016	201624	2016	26	261160	1497.0	

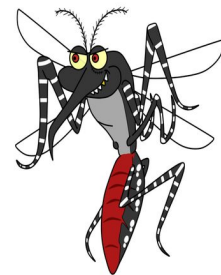
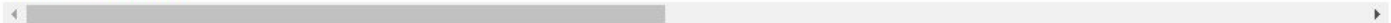
5 rows x 121 columns



'drp df'

	dt_nascimento	tpsexo	tp_gestante	tp_raca_cor	tp_escolaridade	no_bairro_residencia	febre	mialgia	cefaleia	exantema	...	Tp_result_NS1	tp_resu
7313	10-10-1970	F	6.0	9.0	9.0	COHAB	NaN	NaN	NaN	NaN	...	4.0	
11622	23-06-1965	M	6.0	9.0	9.0	IMBIRIBEIRA	NaN	NaN	NaN	NaN	...	4.0	
14426	09-01-2004	M	6.0	9.0	NaN	AREIAS	1.0	1.0	2.0	1.0	...	NaN	
10798	26-02-1964	F	5.0	4.0	8.0	CASA AMARELA	NaN	NaN	NaN	NaN	...	4.0	
15363	03-05-1982	M	6.0	9.0	NaN	BOA VISTA	1.0	1.0	1.0	2.0	...	4.0	

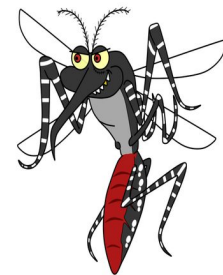
5 rows x 38 columns



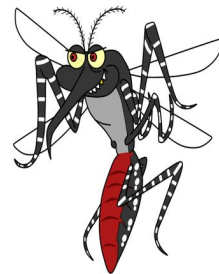
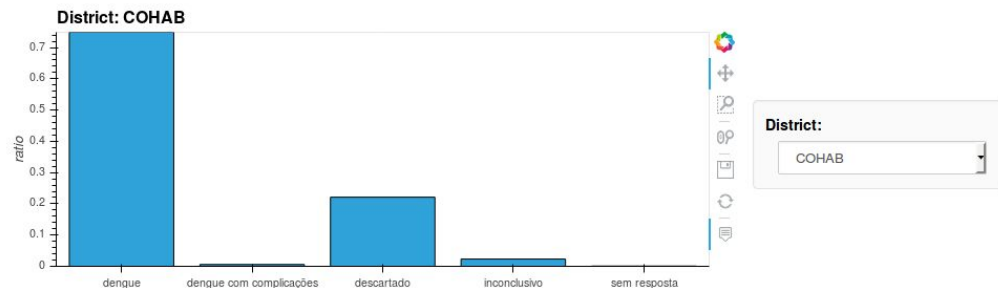
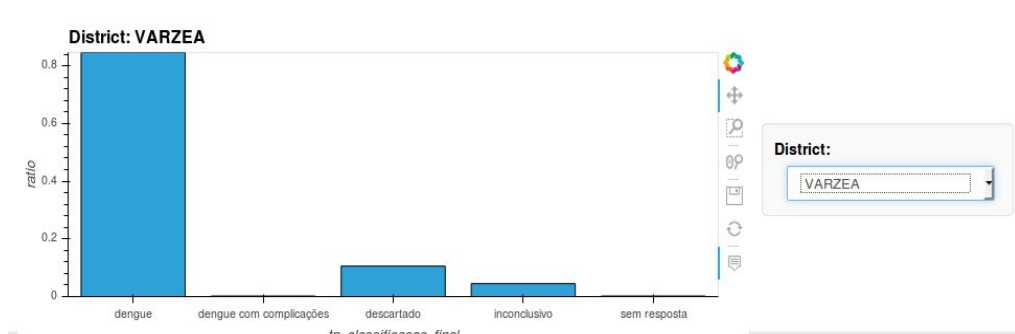
Pré Processamento dos dados

	dt_nascimento	tp_sexo	tp_gestante	tp_raca_cor	tp_escolaridade	no_bairro_residencia	febre	mialgia	cefaleia	exantema	...	Tp_result_NS1	tp_
13827	14-02-1967	F	não se aplica	ignorado	ignorado	JARDIM SAO PAULO	sim	sim	sim	sim	...	sem resposta	
12281	09-07-1958	F	não se aplica	ignorado	ignorado	VARZEA	sem resposta	sem resposta	sem resposta	sem resposta	...	sem resposta	
9319	12-05-1992	F	não se aplica	ignorado	ignorado	CORDEIRO	sim	sim	sim	não	...	sem resposta	
10941	02-08-1974	M	ignorado	ignorado	ignorado	SAN MARTIN	sim	não	não	não	...	sem resposta	
2058	09-03-2008	F	ignorado	ignorado	ignorado	MACAXEIRA	sem resposta	sem resposta	sem resposta	sem resposta	...	não realizado	

5 rows x 38 columns



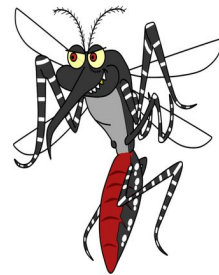
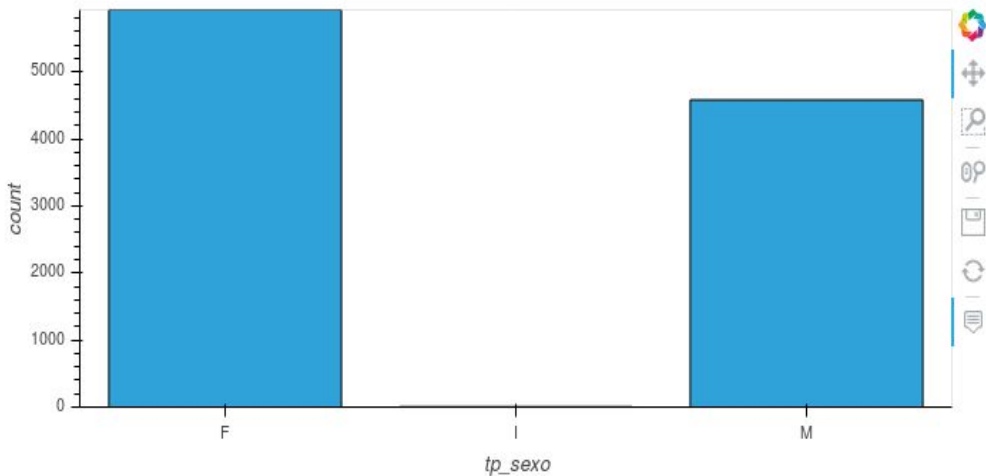
Análise dos tipos de casos registrados por bairro



Análise da quantidade de casos registrados por sexo

tp_sexo	F	I	M
count	5924	6	4581

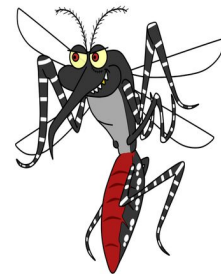
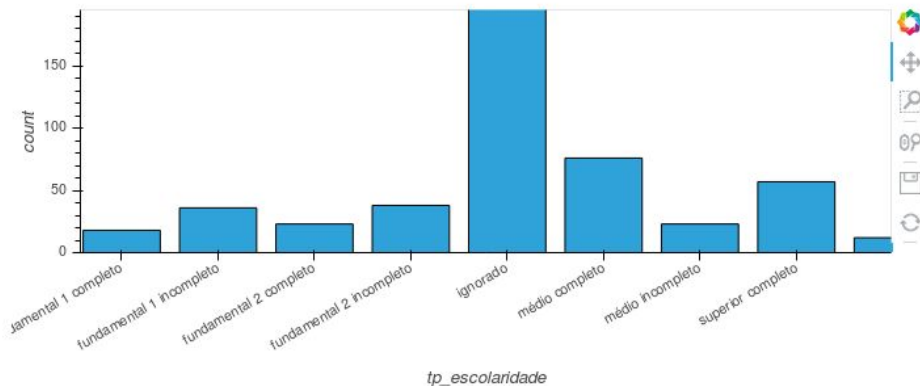
Out[6]:



Análise da quantidade de casos registrados por escolaridade

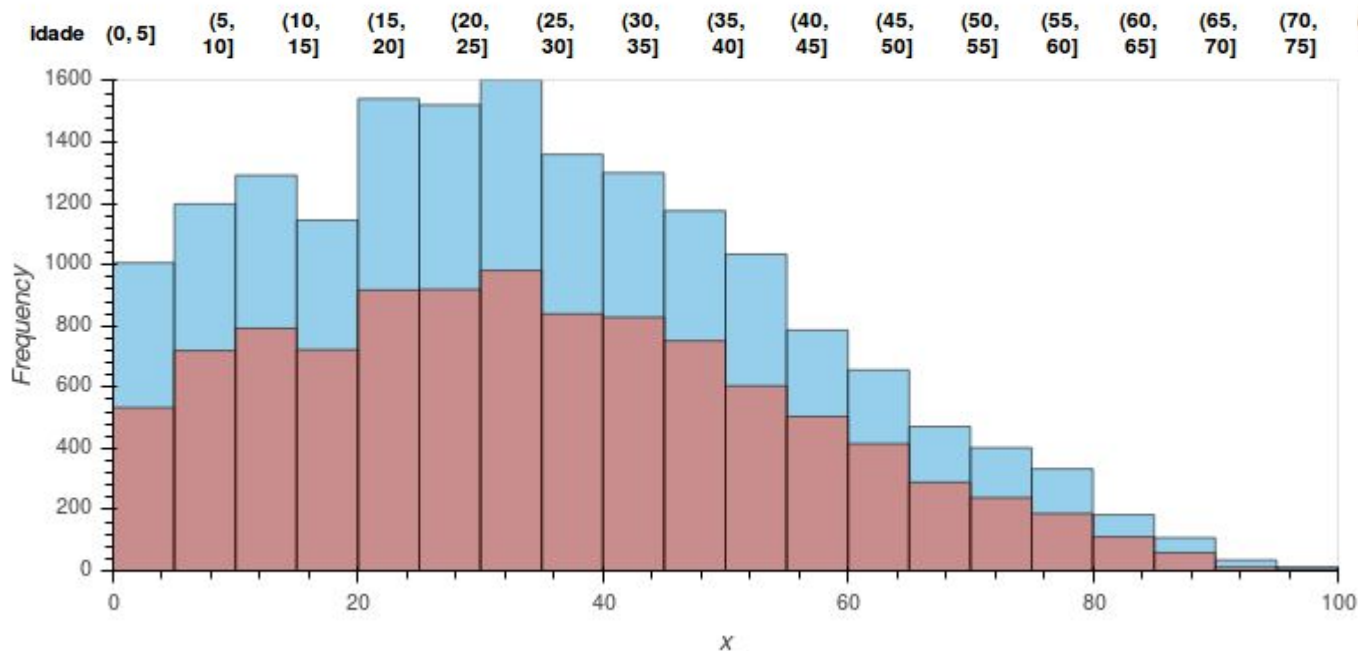
tp_escolaridade	fundamental 1 completo	fundamental 1 Incompleto	fundamental 2 completo	fundamental 2 Incompleto	Ignorado	médio completo	médio Incompleto	superior completo	superior Incompleto
count	18	36	23	38	10228	76	23	57	12

Out[7]:



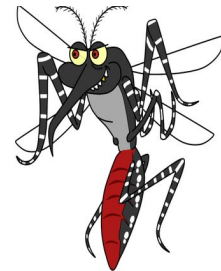
Análise da quantidade de casos registrados por idade

'idade todos'

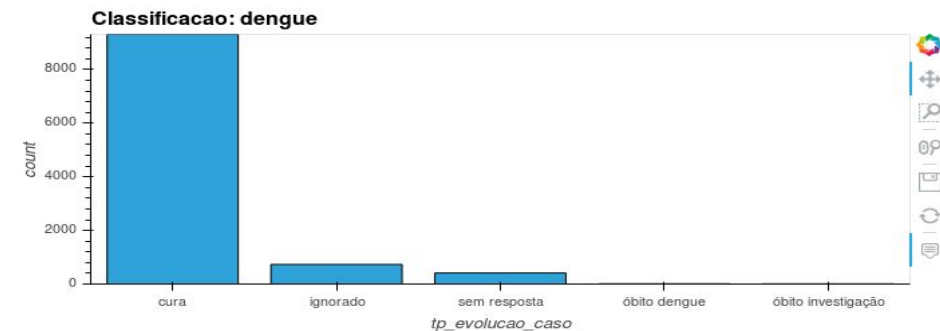


183	108	35	13
-----	-----	----	----

(80, 85]	(85, 90]	(90, 95]	(95, 100]
111	59	12	5

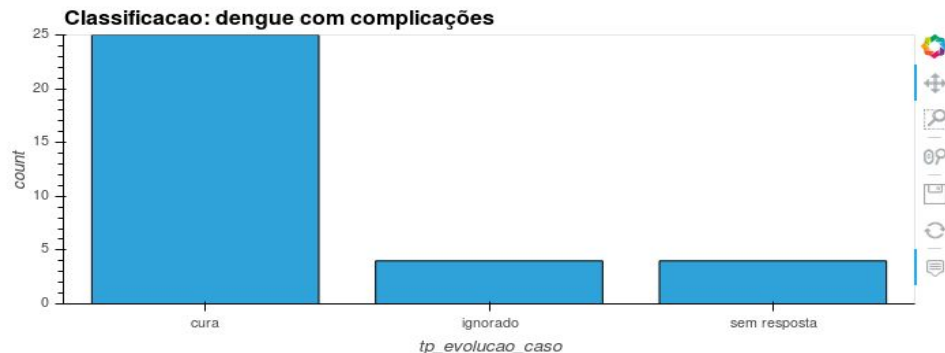


Resultado dos casos por categoria



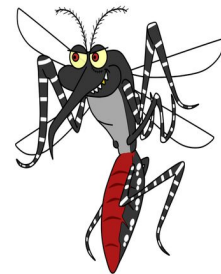
Classificacao:

dengue

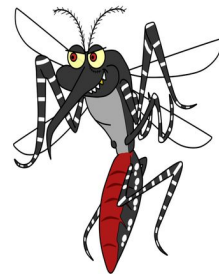
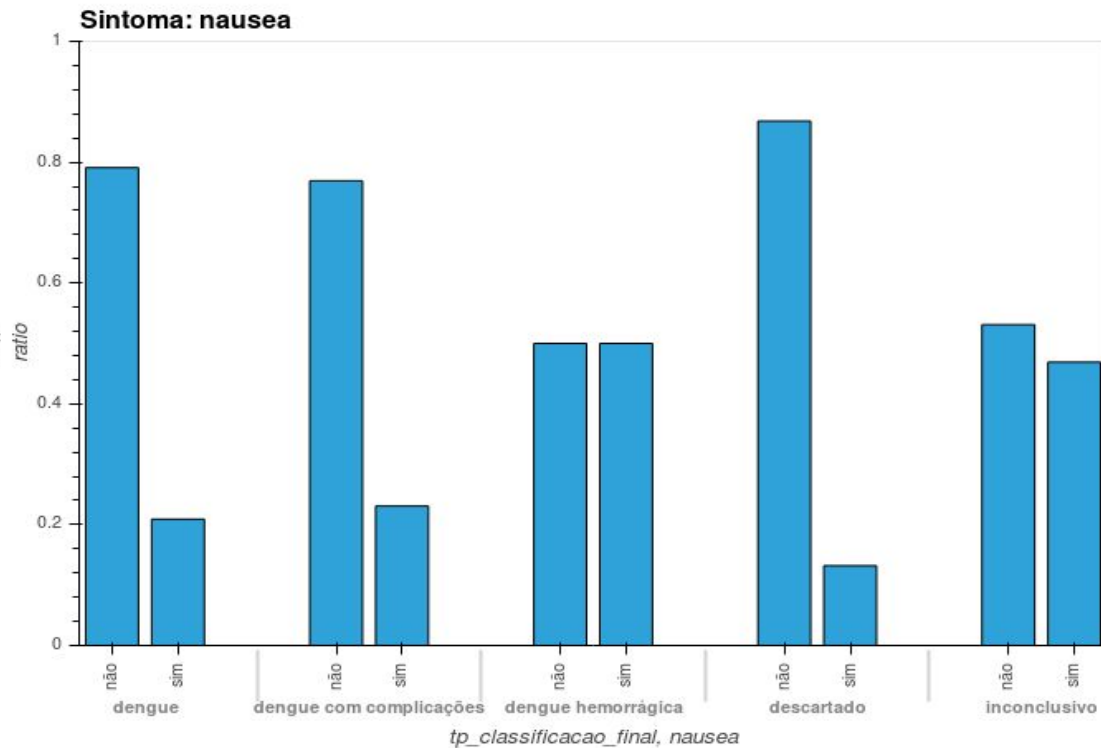


Classificacao:

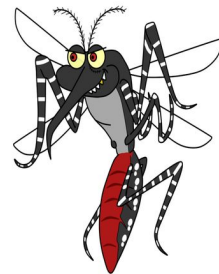
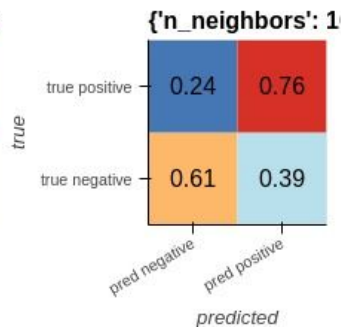
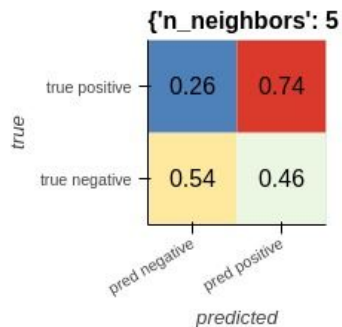
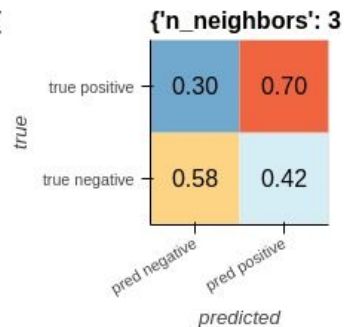
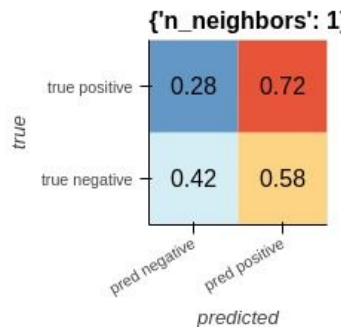
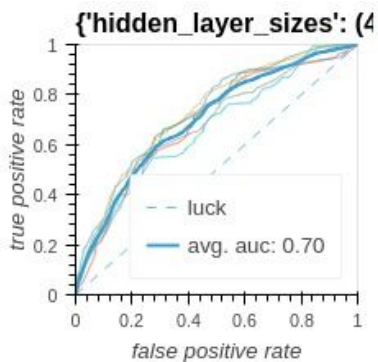
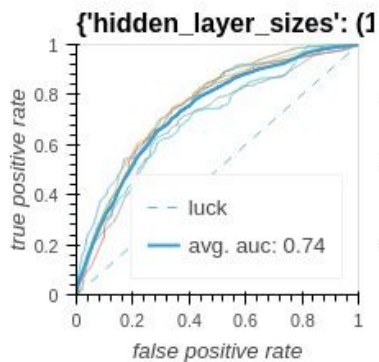
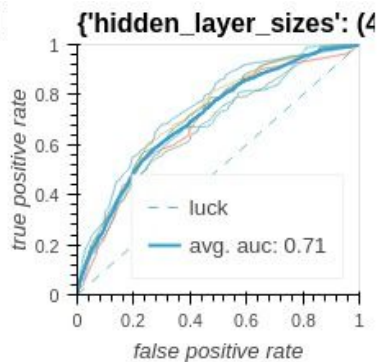
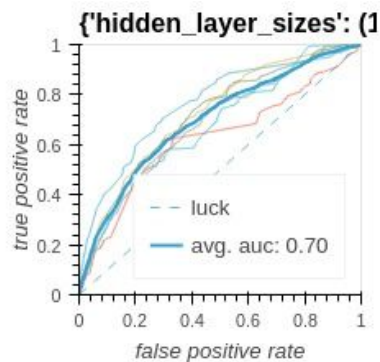
dengue com complicações



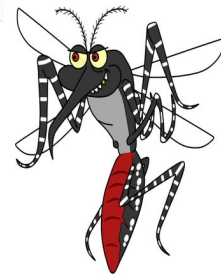
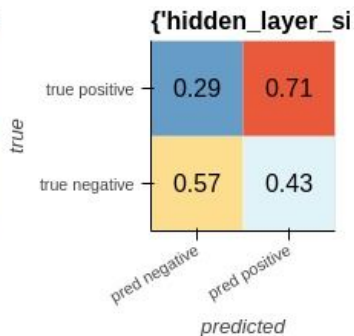
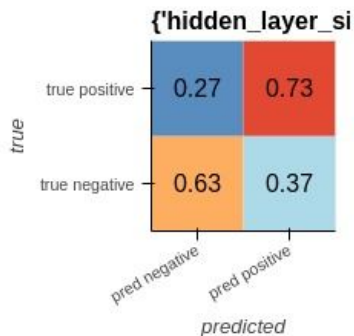
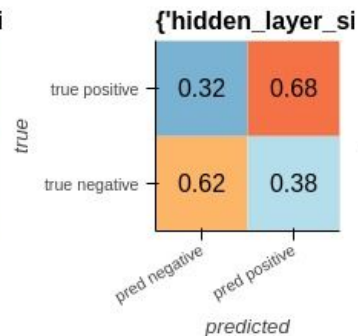
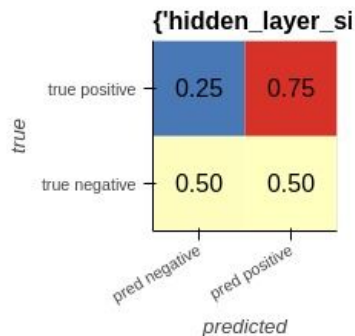
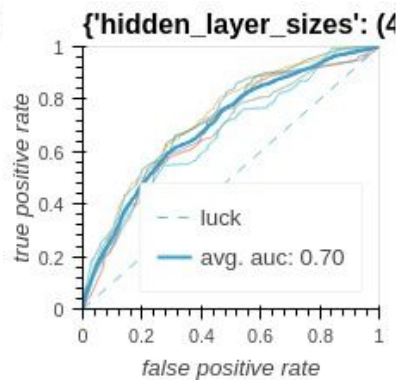
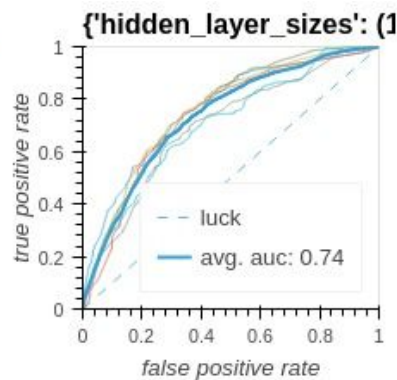
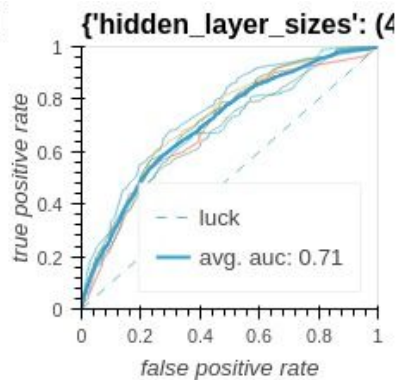
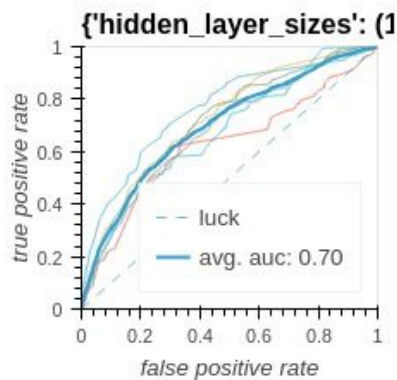
Relação Sintoma X Categoria



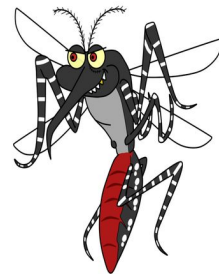
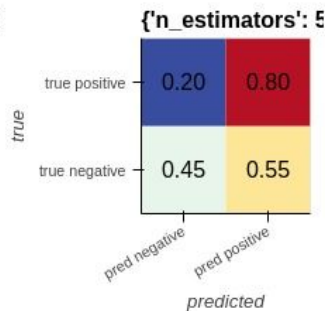
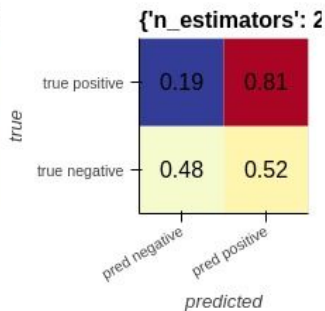
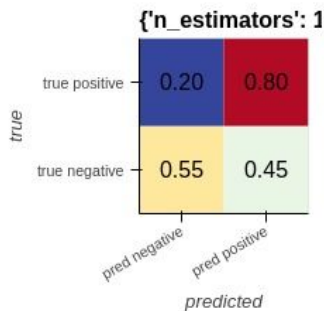
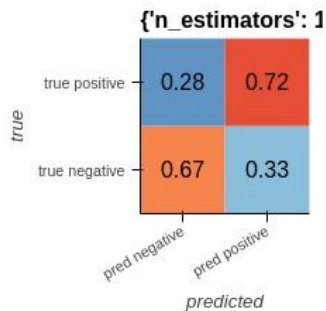
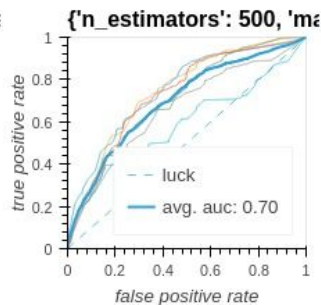
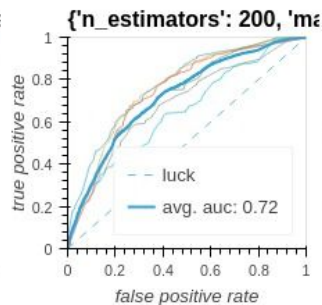
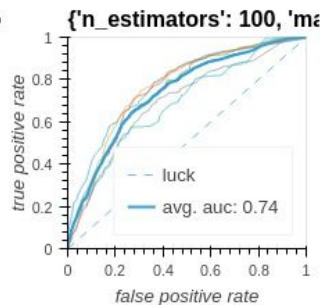
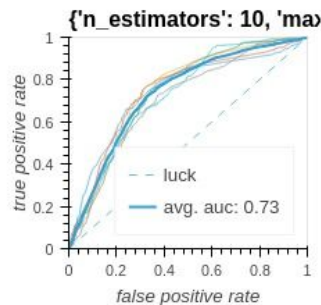
Curva ROC e Matriz de Confusão - KNN



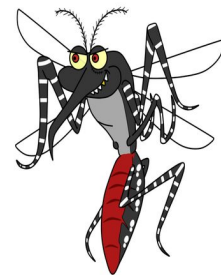
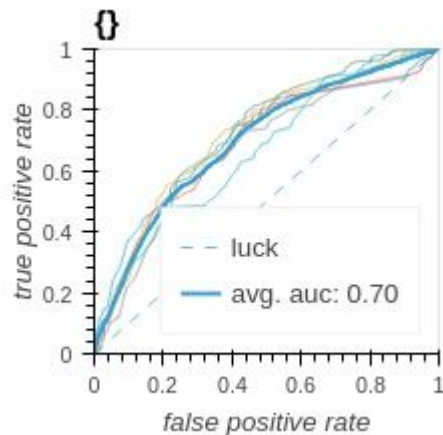
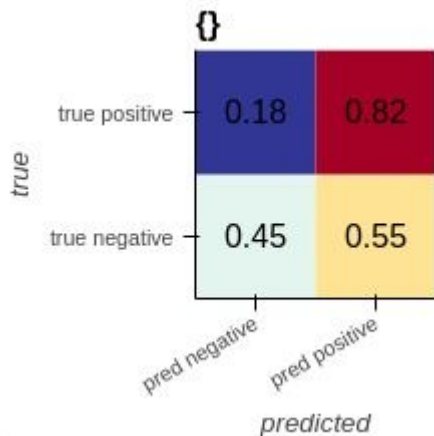
Curva ROC e Matriz de Confusão - MLP



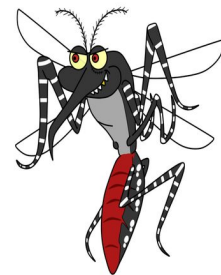
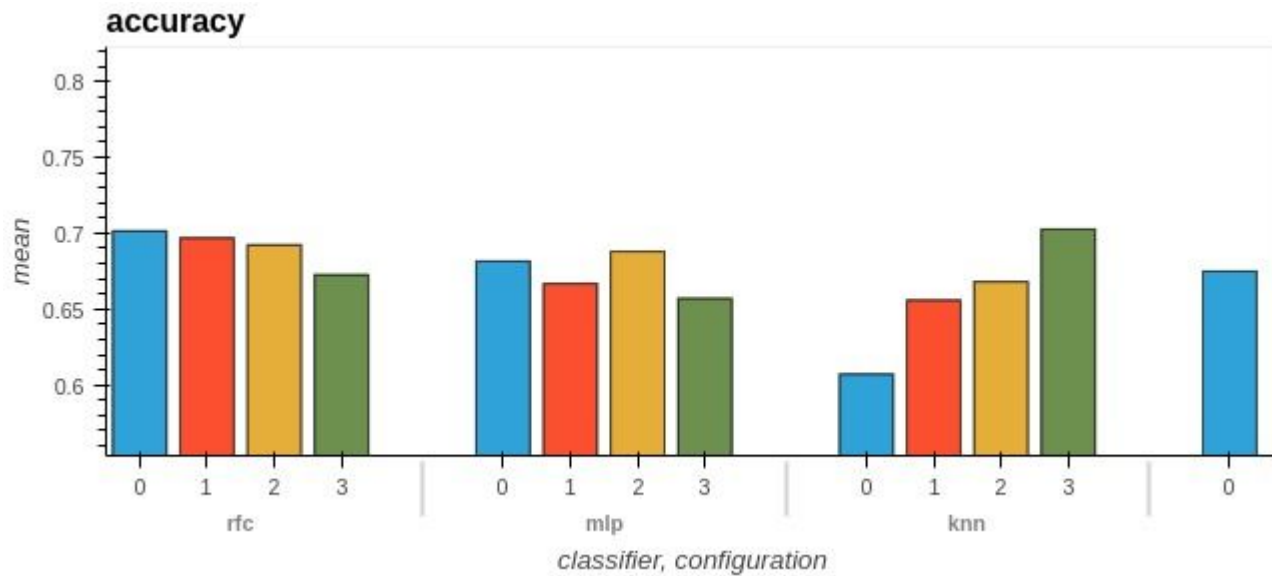
Curva ROC e Matriz de Confusão - Random Forest



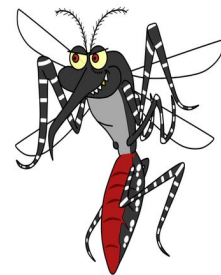
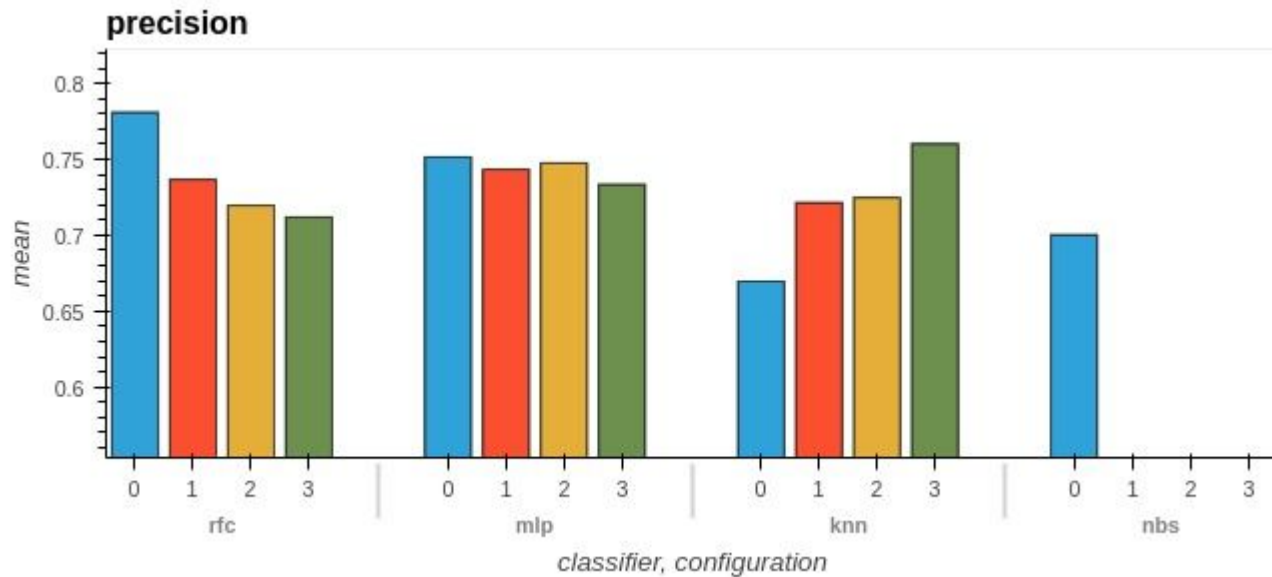
Curva ROC e Matriz de Confusão - Gaussian Bayes



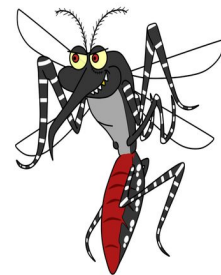
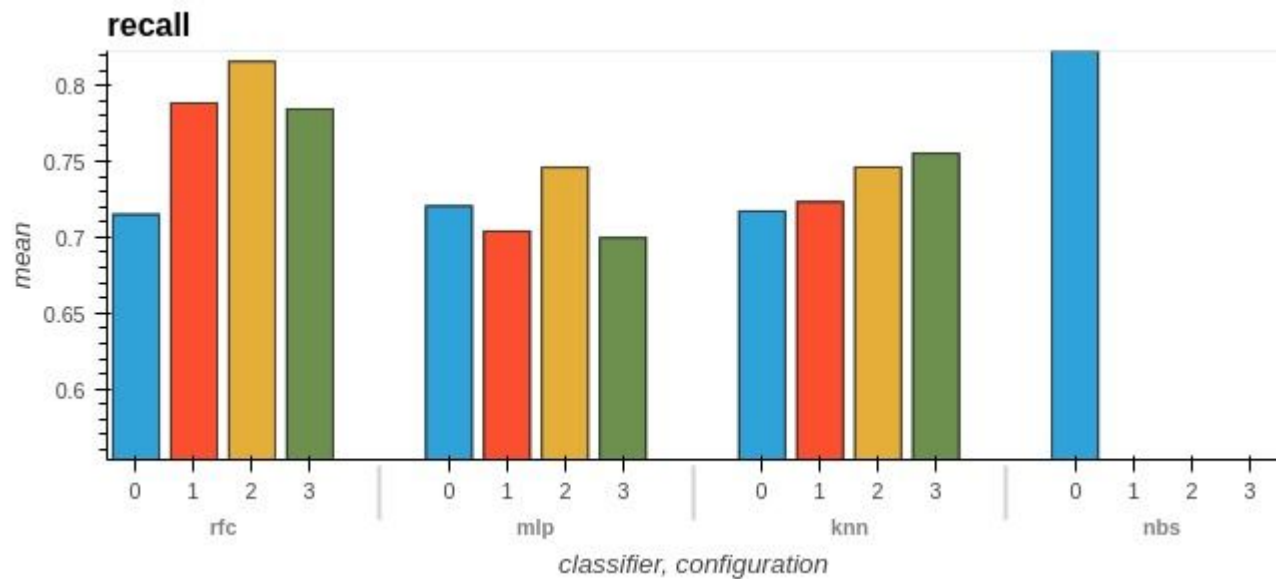
Acurácia



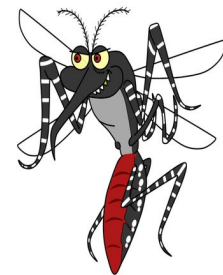
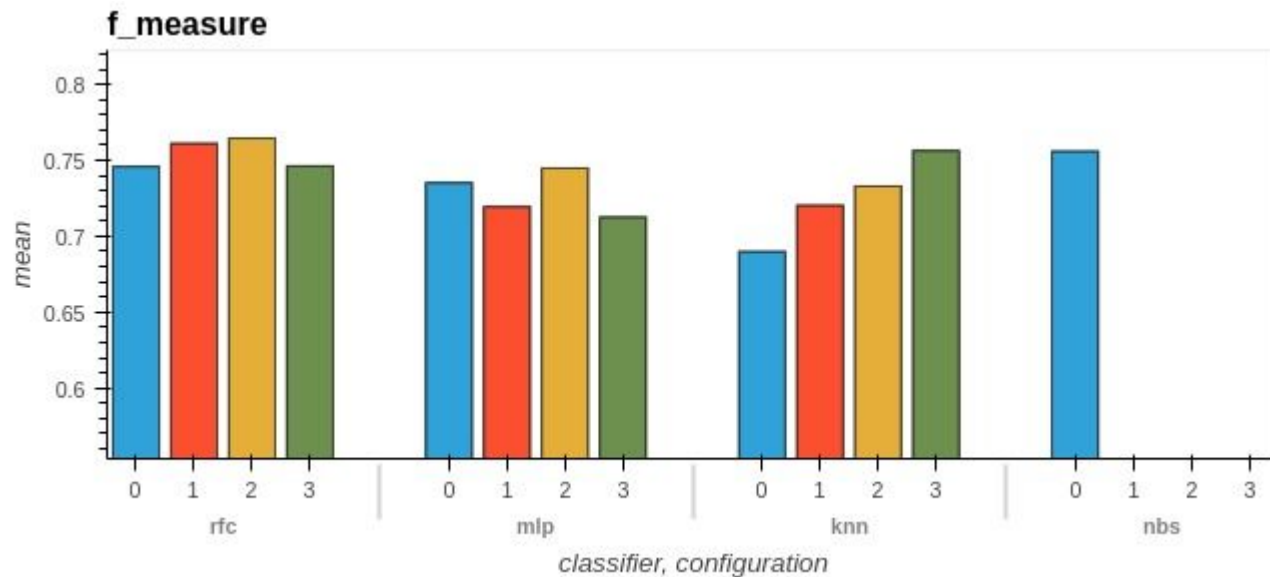
Precisão



Recall



F-Measure



Obrigado!

