

# Appendix A

## Google Play Results

### A.1 Normal

#### A.1.1 No SMOTE and no parameters

	TfidfVectorizer					CountVectorizer					Length				
	P	R	F <sup>1</sup>	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.71	0.81	0.76	0.62	0.48	0.70	0.76	0.73	0.59	0.44	0.74	0.67	0.70	0.59	0.53
RF	0.72	1.00	0.84	0.72	0.70	0.68	0.97	0.80	0.67	0.64	0.67	1.00	0.81	0.59	0.53
SVM	0.72	1.00	0.84	0.72	0.27	0.72	0.97	0.82	0.72	0.33	0.67	1.00	0.81	0.67	0.32
NB	0.75	1.00	0.86	0.76	0.56	0.67	1.00	0.81	0.67	0.50	0.67	1.00	0.81	0.67	0.50
MNB	0.75	1.00	0.86	0.76	0.56	0.83	0.90	0.86	0.70	0.40	0.73	0.90	0.81	0.69	0.64

	Word2Vec					FastText					SBERT				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.68	0.71	0.70	0.55	0.42	0.68	0.75	0.71	0.60	0.54	0.78	0.86	0.82	0.72	0.62
RF	0.70	0.90	0.79	0.63	0.51	0.65	0.86	0.74	0.60	0.56	0.77	0.95	0.85	0.76	0.69
SVM	0.67	1.00	0.81	0.67	0.27	0.65	1.00	0.79	0.65	0.67	0.72	1.00	0.84	0.72	0.25
NB	0.70	0.24	0.36	0.42	0.54	0.83	0.36	0.50	0.53	0.70	0.82	0.86	0.84	0.76	0.73

	USE					Trigrams & CountVectorizer					Trigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.75	0.71	0.73	0.62	0.54	0.72	0.86	0.78	0.66	0.49	0.68	0.71	0.70	0.55	0.42
RF	0.73	0.90	0.81	0.69	0.82	0.68	0.97	0.80	0.67	0.77	0.67	1.00	0.81	0.67	0.76
SVM	0.76	0.90	0.83	0.72	0.27	0.72	0.95	0.81	0.72	0.82	0.72	0.93	0.82	0.72	0.85
NB	0.67	1.00	0.81	0.67	0.50	0.67	1.00	0.81	0.67	0.50	0.67	1.00	0.81	0.67	0.50
MNB	NA	NA	NA	NA	NA	0.75	1.00	0.86	0.76	0.56	0.72	1.00	0.84	0.72	0.53

	Bigrams & CountVectorizer					Bigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.72	0.86	0.78	0.66	0.49	0.68	0.71	0.70	0.55	0.42
RF	0.68	0.97	0.80	0.67	0.66	0.67	1.00	0.81	0.67	0.77
SVM	0.72	0.97	0.82	0.72	0.83	0.73	0.93	0.82	0.72	0.85
NB	0.67	1.00	0.81	0.67	0.50	0.67	1.00	0.81	0.67	0.50
MNB	0.67	1.00	0.81	0.67	0.56	0.72	1.00	0.84	0.72	0.54

<sup>1</sup>F-measure

### A.1.2 SMOTE but no parameters

	TfidfVectorizer					CountVectorizer					Length				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.76	0.76	0.76	0.66	0.57	0.74	0.86	0.79	0.70	0.61	0.67	0.48	0.56	0.45	0.43
RF	0.75	1.00	0.86	0.76	0.68	0.90	0.90	0.90	0.86	0.90	0.75	0.71	0.73	0.62	0.59
SVM	0.71	0.95	0.82	0.69	0.31	0.74	0.86	0.79	0.69	0.31	0.71	0.52	0.60	0.53	0.55
NB	0.75	1.00	0.86	0.76	0.56	0.67	1.00	0.81	0.67	0.50	0.67	1.00	0.81	0.67	0.50
MNB	0.70	0.76	0.73	0.59	0.61	0.75	0.86	0.80	0.69	0.77	0.73	0.90	0.81	0.69	0.58

	Word2Vec					FastText					SBERT				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.85	0.52	0.65	0.56	0.50	0.71	0.57	0.63	0.52	0.47	0.85	0.79	0.82	0.77	0.75
RF	0.67	0.57	0.62	0.48	0.43	0.77	0.71	0.74	0.67	0.63	0.83	0.90	0.86	0.79	0.75
SVM	0.67	0.76	0.71	0.58	0.50	0.42	0.29	0.34	0.28	0.25	0.71	0.95	0.82	0.69	0.28
NB	0.80	0.19	0.31	0.38	0.54	1.00	0.14	0.25	0.38	0.67	0.79	0.90	0.84	0.76	0.70

	USE					Trigrams & CountVectorizer					Trigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.74	0.67	0.70	0.59	0.52	0.68	0.97	0.80	0.67	0.52	0.75	0.86	0.80	0.69	0.55
RF	0.78	0.86	0.82	0.72	0.79	0.72	0.97	0.82	0.72	0.78	0.68	0.97	0.80	0.67	0.71
SVM	0.71	0.95	0.82	0.69	0.29	0.74	1.00	0.85	0.76	0.77	0.67	1.00	0.81	0.67	0.72
NB	0.78	0.86	0.82	0.72	0.72	0.67	1.00	0.81	0.67	0.50	0.67	1.00	0.81	0.66	0.56
MNB	NA	NA	NA	NA	NA	0.75	1.00	0.86	0.76	0.56	0.67	1.00	0.81	0.66	0.56

	Bigrams & CountVectorizer					Bigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.68	0.97	0.80	0.67	0.52	0.74	0.81	0.77	0.66	0.53
RF	0.72	0.97	0.82	0.72	0.74	0.68	0.97	0.80	0.67	0.74
SVM	0.74	1.00	0.85	0.76	0.78	0.67	1.00	0.81	0.67	0.85
NB	0.67	1.00	0.81	0.67	0.50	0.67	1.00	0.81	0.67	0.56
MNB	0.67	1.00	0.81	0.67	0.50	0.72	1.00	0.84	0.72	0.54

### A.1.3 SMOTE + parameters (gridsearch)

	TfidfVectorizer					CountVectorizer					Length				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.78	0.67	0.72	0.62	0.56	0.83	0.95	0.89	0.83	0.74	0.79	0.52	0.63	0.55	0.56
RF	0.75	1.00	0.86	0.76	0.68	0.84	0.76	0.80	0.72	0.78	0.75	0.71	0.73	0.62	0.60
SVM	0.74	0.95	0.83	0.72	0.29	1.00	0.24	0.38	0.45	0.82	0.75	0.43	0.55	0.48	0.53
NB	0.89	0.76	0.82	0.76	0.75	0.83	0.71	0.77	0.69	0.70	0.85	0.81	0.83	0.76	0.65
MNB	0.73	0.90	0.81	0.69	0.65	0.73	0.90	0.81	0.69	0.73	0.73	0.90	0.81	0.69	0.58

	Word2Vec					FastText					SBERT				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.80	0.76	0.78	0.69	0.62	0.71	0.57	0.63	0.52	0.44	0.77	0.81	0.79	0.69	0.59
RF	0.67	0.57	0.62	0.48	0.41	0.71	0.81	0.76	0.62	0.43	0.83	0.90	0.86	0.79	0.77
SVM	0.67	0.76	0.71	0.69	0.32	0.42	0.29	0.34	0.28	0.30	0.67	0.10	0.17	0.31	0.30
NB	0.80	0.19	0.31	0.38	0.63	1.00	0.14	0.25	0.38	0.67	<b>0.79</b>	<b>0.90</b>	<b>0.84</b>	<b>0.76</b>	<b>0.70</b>

	USE					Trigrams & CountVectorizer					Trigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.71	0.71	0.71	0.59	0.49	0.68	0.97	0.80	0.67	0.59	0.73	0.90	0.81	0.69	0.49
RF	0.78	0.86	0.82	0.72	0.80	0.72	0.97	0.82	0.72	0.78	0.75	1.00	0.86	0.76	0.70
SVM	0.67	0.10	0.17	0.31	0.35	0.74	1.00	0.85	0.76	0.77	0.67	0.10	0.17	0.31	0.35
NB	0.74	0.90	0.81	0.72	0.68	0.67	1.00	0.81	0.67	0.50	1.00	0.05	0.09	0.31	0.77
MNB	NA	NA	NA	NA	NA	0.75	1.00	0.86	0.76	0.56	0.78	0.86	0.82	0.72	0.78

	Bigrams & CountVectorizer					Bigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.68	0.97	0.80	0.67	0.52	0.73	0.90	0.81	0.69	0.49
RF	0.72	0.97	0.82	0.72	0.74	0.68	0.97	0.80	0.67	0.74
SVM	0.74	1.00	0.85	0.77	0.78	0.67	1.00	0.81	0.67	0.85
NB	0.67	1.00	0.81	0.67	0.50	1.00	0.05	0.09	0.31	0.77
MNB	0.67	1.00	0.81	0.67	0.56	0.75	0.86	0.80	0.69	0.77

## A.2 Stemming

### A.2.1 No SMOTE and no parameters

	TfidfVectorizer					CountVectorizer					Length				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.73	0.90	0.81	0.70	0.60	0.76	0.76	0.76	0.66	0.42	0.76	0.76	0.76	0.66	0.57
RF	0.72	1.00	0.84	0.72	0.73	0.72	1.00	0.84	0.72	0.78	0.75	1.00	0.86	0.76	0.66
SVM	0.72	1.00	0.84	0.72	0.85	0.72	1.00	0.84	0.72	0.85	0.67	1.00	0.81	0.67	0.85
NB	0.77	0.95	0.85	0.76	0.60	0.75	1.00	0.86	0.76	0.56	0.77	0.95	0.85	0.76	0.60
MNB	0.72	1.00	0.84	0.72	0.54	0.72	1.00	0.84	0.72	0.54	0.67	1.00	0.81	0.67	0.50

	Word2Vec					FastText					SBERT				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.81	0.62	0.70	0.62	0.51	0.67	0.57	0.62	0.48	0.41	0.74	0.67	0.70	0.59	0.52
RF	0.72	1.00	0.84	0.72	0.51	0.77	0.81	0.79	0.69	0.54	0.73	0.90	0.81	0.69	0.58
SVM	0.72	1.00	0.84	0.72	0.35	0.72	1.00	0.84	0.72	0.34	0.71	0.95	0.82	0.69	0.66
NB	0.50	0.14	0.22	0.28	0.22	0.75	1.00	0.86	0.76	0.55	0.70	0.66	0.68	0.58	0.61

	USE					Trigrams & CountVectorizer					Trigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.72	0.86	0.78	0.66	0.65	0.73	0.76	0.74	0.62	0.51	0.73	0.93	0.82	0.72	0.61
RF	0.76	0.90	0.83	0.69	0.76	0.75	1.00	0.86	0.76	0.63	0.72	1.00	0.84	0.72	0.75
SVM	0.79	0.90	0.84	0.77	0.80	0.72	0.97	0.82	0.72	0.85	0.72	1.00	0.84	0.72	0.83
NB	0.78	1.00	0.88	0.76	0.78	0.67	1.00	0.81	0.67	0.63	0.72	1.00	0.84	0.72	0.63
MNB	NA	NA	NA	NA	NA	0.67	1.00	0.81	0.67	0.79	0.72	1.00	0.84	0.72	0.53

	Bigrams & CountVectorizer					Bigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.73	0.90	0.81	0.69	0.51	0.79	0.90	0.84	0.76	0.64
RF	0.75	1.00	0.86	0.76	0.59	0.72	1.00	0.84	0.72	0.76
SVM	0.72	0.97	0.82	0.72	0.85	0.72	1.00	0.84	0.72	0.84
NB	0.77	0.95	0.85	0.76	0.60	0.77	0.95	0.85	0.76	0.60
MNB	0.67	1.00	0.81	0.67	0.50	0.72	1.00	0.84	0.72	0.53

### A.2.2 SMOTE but no parameters

	TfidfVectorizer					CountVectorizer					Length				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.71	0.81	0.76	0.62	0.47	0.74	0.81	0.77	0.66	0.53	0.73	0.76	0.74	0.62	0.51
RF	0.75	1.00	0.86	0.76	0.70	0.82	0.79	0.81	0.74	0.80	0.81	0.76	0.79	0.72	0.82
SVM	0.83	0.90	0.86	0.79	0.83	0.72	0.97	0.82	0.72	0.81	0.73	0.52	0.61	0.52	0.55
NB	0.77	0.95	0.85	0.76	0.60	0.75	1.00	0.86	0.76	0.56	0.75	1.00	0.86	0.76	0.56
MNB	0.89	0.76	0.82	0.76	0.60	0.67	1.00	0.81	0.67	0.50	0.67	1.00	0.81	0.67	0.50

	Word2Vec					FastText					SBERT				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.68	0.62	0.65	0.52	0.43	0.73	0.52	0.61	0.52	0.51	0.79	0.71	0.75	0.66	0.51
RF	0.73	0.90	0.81	0.69	0.37	0.64	0.43	0.51	0.41	0.42	0.75	0.86	0.80	0.69	0.63
SVM	0.68	0.62	0.65	0.52	0.42	0.68	0.62	0.65	0.52	0.24	0.75	0.86	0.80	0.69	0.72
NB	0.43	0.14	0.21	0.24	0.19	0.78	0.67	0.72	0.62	0.62	0.78	0.86	0.82	0.72	0.65

	USE					Trigrams & CountVectorizer					Trigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.73	0.76	0.74	0.62	0.51	0.81	0.81	0.81	0.72	0.65	0.71	0.71	0.71	0.59	0.47
RF	0.75	0.86	0.80	0.69	0.74	0.72	0.97	0.82	0.72	0.71	0.68	0.97	0.80	0.67	0.73
SVM	0.74	0.97	0.84	0.74	0.80	0.72	1.00	0.84	0.72	0.85	0.67	1.00	0.81	0.67	0.73
NB	0.75	0.86	0.80	0.69	0.67	0.67	1.00	0.81	0.67	0.63	0.69	1.00	0.82	0.70	0.53
MNB	NA	NA	NA	NA	NA	0.67	1.00	0.81	0.67	0.63	0.88	0.71	0.79	0.72	0.77

	Bigrams & CountVectorizer					Bigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.74	0.81	0.77	0.66	0.58	0.68	0.71	0.70	0.55	0.47
RF	0.72	1.00	0.84	0.72	0.78	0.68	0.97	0.80	0.67	0.75
SVM	0.72	0.97	0.82	0.72	0.85	0.75	0.93	0.83	0.74	0.82
NB	0.77	0.95	0.85	0.76	0.60	0.77	0.95	0.85	0.76	0.60
MNB	0.67	1.00	0.81	0.67	0.50	0.89	0.76	0.82	0.70	0.54

### A.2.3 SMOTE + parameters - grid

	TfidfVectorizer					CountVectorizer					Length				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.75	0.86	0.80	0.69	0.55	0.73	0.90	0.81	0.69	0.49	0.73	0.90	0.81	0.69	0.54
RF	0.75	1.00	0.86	0.76	0.71	0.82	0.79	0.81	0.74	0.80	0.84	0.76	0.80	0.72	0.81
SVM	0.78	1.00	0.88	0.79	0.86	0.74	0.86	0.79	0.70	0.72	0.71	0.52	0.60	0.53	0.55
NB	1.00	0.24	0.38	0.45	0.74	1.00	0.05	0.09	0.31	0.77	0.78	0.33	0.47	0.45	0.72
MNB	0.82	0.86	0.84	0.76	0.75	0.75	0.86	0.80	0.69	0.77	0.67	1.00	0.81	0.67	0.50

	Word2Vec					FastText					SBERT				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.68	0.62	0.65	0.52	0.54	0.71	0.57	0.63	0.52	0.52	0.78	0.67	0.72	0.62	0.69
RF	0.71	0.81	0.76	0.62	0.42	0.73	0.52	0.61	0.52	0.52	0.75	0.86	0.80	0.69	0.63
SVM	0.72	0.62	0.67	0.55	0.45	0.68	0.62	0.65	0.52	0.24	0.76	0.90	0.83	0.72	0.70
NB	0.43	0.14	0.21	0.24	0.19	0.78	0.67	0.72	0.62	0.62	0.77	0.81	0.79	0.69	0.68

	USE					Trigrams & CountVectorizer					Trigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.76	0.90	0.83	0.72	0.58	0.81	0.81	0.81	0.72	0.64	0.71	0.81	0.76	0.62	0.60
RF	0.78	0.86	0.82	0.72	0.74	0.72	0.97	0.82	0.72	0.71	0.68	0.97	0.80	0.67	0.73
SVM	0.74	0.97	0.84	0.74	0.80	0.76	1.00	0.87	0.79	0.78	0.67	1.00	0.81	0.67	0.73
NB	0.89	0.81	0.85	0.74	0.77	0.67	1.00	0.81	0.67	0.63	1.00	0.19	0.32	0.41	0.71
MNB	NA	NA	NA	NA	NA	0.67	1.00	0.81	0.67	0.63	0.69	1.00	0.82	0.70	0.53

	Bigrams & CountVectorizer					Bigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.81	0.81	0.81	0.72	0.63	0.81	0.81	0.81	0.72	0.63
RF	0.72	0.97	0.82	0.72	0.81	0.68	0.97	0.80	0.67	0.75
SVM	0.76	1.00	0.87	0.79	0.78	0.67	1.00	0.81	0.67	0.86
NB	0.81	0.81	0.81	0.72	0.63	1.00	0.33	0.50	0.52	0.74
MNB	0.67	1.00	0.81	0.67	0.50	0.89	0.76	0.82	0.70	0.54

### A.3 Lemmatization

#### A.3.1 No SMOTE and no parameters

	TfidfVectorizer					CountVectorizer					Length				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.77	0.95	0.85	0.70	0.60	0.68	0.71	0.70	0.55	0.42	0.76	0.76	0.76	0.66	0.57
RF	0.72	1.00	0.84	0.72	0.73	0.72	1.00	0.84	0.72	0.78	0.75	1.00	0.86	0.76	0.66
SVM	0.72	1.00	0.84	0.72	0.85	0.72	1.00	0.84	0.72	0.85	0.67	1.00	0.81	0.67	0.85
NB	0.77	0.95	0.85	0.76	0.60	0.75	1.00	0.86	0.76	0.56	0.75	1.00	0.86	0.76	0.56
MNB	0.72	1.00	0.84	0.72	0.54	0.72	1.00	0.84	0.72	0.54	0.67	1.00	0.81	0.67	0.50

	Word2Vec					FastText					SBERT				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.73	0.76	0.74	0.62	0.51	0.67	0.57	0.62	0.48	0.41	0.74	0.67	0.70	0.59	0.52
RF	0.72	1.00	0.84	0.72	0.51	0.76	0.90	0.83	0.72	0.59	0.74	0.95	0.83	0.72	0.58
SVM	0.72	1.00	0.84	0.72	0.35	0.72	1.00	0.84	0.72	0.34	0.71	0.95	0.82	0.69	0.66
NB	0.50	0.14	0.22	0.28	0.22	0.75	1.00	0.86	0.76	0.55	0.70	0.66	0.68	0.58	0.61

	USE					Trigrams & CountVectorizer					Trigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.81	0.81	0.81	0.72	0.65	0.73	0.90	0.81	0.69	0.51	0.73	0.93	0.82	0.72	0.61
RF	0.76	0.90	0.83	0.72	0.76	0.75	1.00	0.86	0.76	0.63	0.72	1.00	0.84	0.72	0.75
SVM	0.79	0.90	0.84	0.77	0.80	0.72	0.97	0.82	0.72	0.85	0.72	1.00	0.84	0.72	0.83
NB	0.82	0.86	0.84	0.76	0.78	0.67	1.00	0.81	0.67	0.63	0.78	1.00	0.88	0.79	0.63
MNB	NA	NA	NA	NA	NA	0.67	1.00	0.81	0.67	0.79	0.72	1.00	0.84	0.72	0.53

	Bigrams & CountVectorizer					Bigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.73	0.90	0.81	0.69	0.51	0.79	0.90	0.84	0.76	0.64
RF	0.75	1.00	0.86	0.76	0.59	0.72	1.00	0.84	0.72	0.76
SVM	0.72	0.97	0.82	0.72	0.85	0.72	1.00	0.84	0.72	0.84
NB	0.77	0.95	0.85	0.76	0.60	0.77	0.95	0.85	0.76	0.60
MNB	0.67	1.00	0.81	0.67	0.50	0.72	1.00	0.84	0.72	0.53

### A.3.2 SMOTE but no parameters

	TfidfVectorizer					CountVectorizer					Length				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.71	0.81	0.76	0.62	0.47	0.74	0.81	0.77	0.66	0.53	0.74	0.81	0.77	0.66	0.53
RF	0.75	1.00	0.86	0.76	0.70	0.82	0.79	0.81	0.74	0.80	0.81	0.76	0.79	0.72	0.82
SVM	0.83	0.90	0.86	0.79	0.83	0.72	0.97	0.82	0.72	0.81	0.76	0.62	0.68	0.59	0.5
NB	0.77	0.95	0.85	0.76	0.60	0.75	1.00	0.86	0.76	0.56	0.75	1.00	0.86	0.76	0.56
MNB	0.77	0.95	0.85	0.76	0.60	0.67	1.00	0.81	0.67	0.50	0.67	1.00	0.81	0.67	0.50

	Word2Vec					FastText					SBERT				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.74	0.67	0.70	0.59	0.52	0.73	0.52	0.61	0.52	0.51	0.73	0.90	0.81	0.69	0.51
RF	0.73	0.90	0.81	0.69	0.37	0.73	0.52	0.61	0.52	0.52	0.75	0.86	0.80	0.69	0.63
SVM	0.68	0.62	0.65	0.52	0.42	0.68	0.62	0.65	0.52	0.24	0.76	0.90	0.83	0.72	0.72
NB	0.43	0.14	0.21	0.24	0.19	0.78	0.67	0.72	0.62	0.62	0.78	0.86	0.82	0.72	0.65

	USE					Trigrams & CountVectorizer					Trigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.75	0.86	0.80	0.69	0.55	0.81	0.81	0.81	0.72	0.65	0.71	0.81	0.76	0.62	0.47
RF	0.78	0.86	0.82	0.72	0.74	0.72	0.97	0.82	0.72	0.71	0.68	0.97	0.80	0.67	0.73
SVM	0.74	0.97	0.84	0.74	0.80	0.72	1.00	0.84	0.72	0.85	0.67	1.00	0.81	0.67	0.73
NB	0.78	0.86	0.82	0.72	0.67	0.67	1.00	0.81	0.67	0.63	0.69	1.00	0.82	0.70	0.53
MNB	NA	NA	NA	NA	NA	0.67	1.00	0.81	0.67	0.63	0.88	0.71	0.79	0.72	0.77

	Bigrams & CountVectorizer					Bigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.76	0.90	0.83	0.72	0.58	0.71	0.81	0.76	0.62	0.47
RF	0.72	1.00	0.84	0.72	0.78	0.68	0.97	0.80	0.67	0.75
SVM	0.72	0.97	0.82	0.72	0.85	0.75	0.93	0.83	0.74	0.82
NB	0.77	0.95	0.85	0.76	0.60	0.77	0.95	0.85	0.76	0.60
MNB	0.67	1.00	0.81	0.67	0.50	0.89	0.76	0.82	0.70	0.54

### A.3.3 SMOTE + parameters - grid

	TfidfVectorizer					CountVectorizer					Length				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.75	0.86	0.80	0.69	0.55	0.73	0.90	0.81	0.69	0.49	0.73	0.90	0.81	0.69	0.54
RF	0.75	1.00	0.86	0.76	0.71	0.82	0.79	0.81	0.74	0.80	0.84	0.76	0.80	0.72	0.81
SVM	0.78	1.00	0.88	0.79	0.86	0.74	0.86	0.79	0.70	0.72	0.71	0.52	0.60	0.53	0.55
NB	1.00	0.24	0.38	0.45	0.74	1.00	0.05	0.09	0.31	0.77	0.78	0.33	0.47	0.45	0.72
MNB	0.82	0.86	0.84	0.76	0.75	0.75	0.86	0.80	0.69	0.77	0.67	1.00	0.81	0.67	0.50

	Word2Vec					FastText					SBERT				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.76	0.76	0.76	0.66	0.54	0.71	0.57	0.63	0.52	0.52	0.80	0.76	0.78	0.69	0.69
RF	0.72	0.86	0.78	0.66	0.38	0.73	0.52	0.61	0.52	0.52	0.75	0.86	0.80	0.69	0.63
SVM	0.72	0.62	0.67	0.55	0.45	0.68	0.62	0.65	0.52	0.24	0.76	0.90	0.83	0.72	0.70
NB	0.43	0.14	0.21	0.24	0.19	0.78	0.67	0.72	0.62	0.62	0.77	0.81	0.79	0.69	0.68

	USE					Trigrams & CountVectorizer					Trigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.76	0.90	0.83	0.72	0.58	0.81	0.81	0.81	0.72	0.64	0.71	0.81	0.76	0.62	0.60
RF	0.78	0.86	0.82	0.72	0.74	0.72	0.97	0.82	0.72	0.71	0.68	0.97	0.80	0.67	0.73
SVM	0.74	0.97	0.84	0.74	0.80	0.76	1.00	0.87	0.79	0.78	0.67	1.00	0.81	0.67	0.73
NB	0.89	0.81	0.85	0.74	0.77	0.67	1.00	0.81	0.67	0.63	1.00	0.19	0.32	0.41	0.71
MNB	NA	NA	NA	NA	NA	0.67	1.00	0.81	0.67	0.63	0.69	1.00	0.82	0.70	0.53

	Bigrams & CountVectorizer					Bigrams & TfidfVectorizer				
	P	R	F	A	AUC	P	R	F	A	AUC
DT	0.77	0.95	0.85	0.76	0.66	0.80	0.76	0.78	0.69	0.66
RF	0.72	0.97	0.82	0.72	0.81	0.68	0.97	0.80	0.67	0.75
SVM	0.76	1.00	0.87	0.79	0.78	0.67	1.00	0.81	0.67	0.86
NB	0.67	1.00	0.81	0.67	0.50	1.00	0.24	0.38	0.45	0.72
MNB	0.67	1.00	0.81	0.67	0.50	0.89	0.76	0.82	0.70	0.54