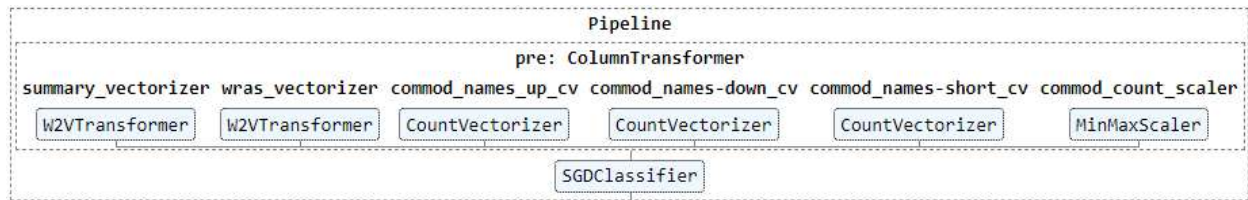
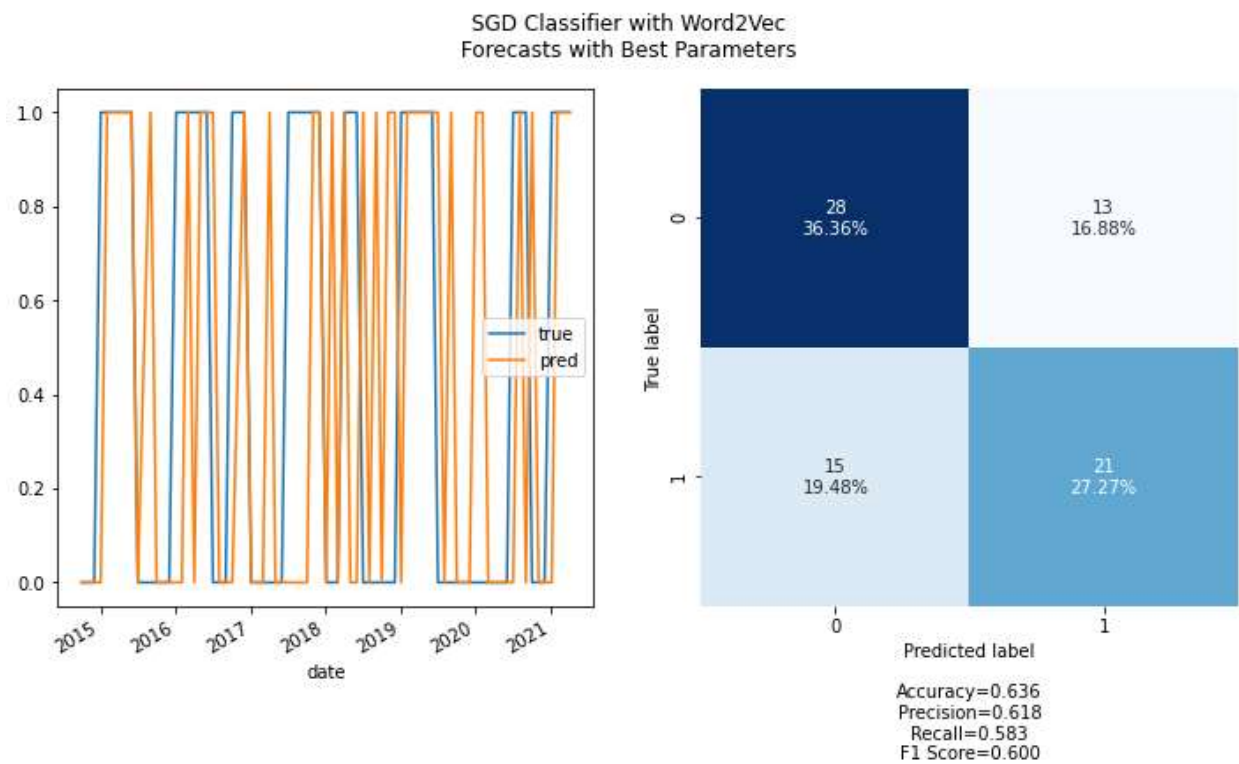


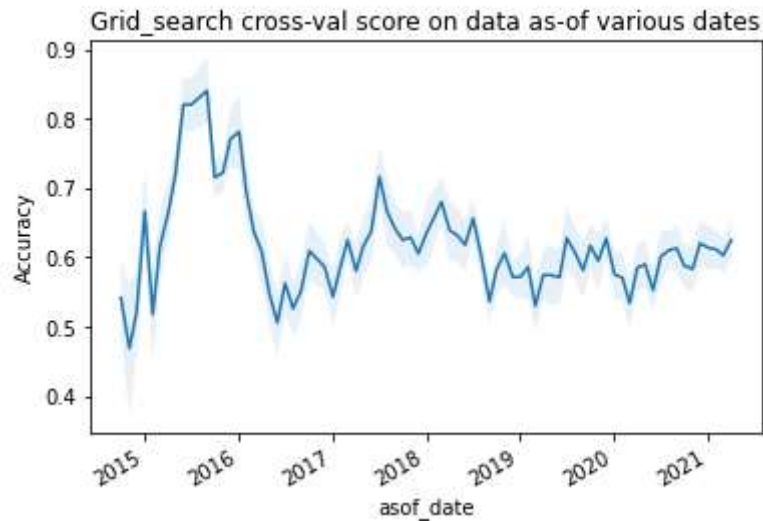
Stochastic Gradient Descent (SGD)

Pipeline



Metrics:





Best Params:

`sgd_w2v_pipeline.get_params()`

```
{'memory': '/tmp/tmp6lgxcgrd',
 'pre': ColumnTransformer(transformers=[('summary_vectorizer',
W2VTransformer(),
                                'clean_summary'),
('wras_vectorizer', W2VTransformer(),
 'clean_wras'),
('commod_names_up_cv',
CountVectorizer(),
 'commodities_up_in_price'),
('commod_names-down_cv',
CountVectorizer(),
 'commodities_down_in_price'),
('commod_names-short_cv',
CountVectorizer(),
 'commodities_in_short_supply'),
('commod_count_scaler',
MinMaxScaler(),
 ['commodities_up_count',
 'commodities_down_count',
 'commodities_short_count'])]),
 'pre__commod_count_scaler': MinMaxScaler(),
 'pre__commod_count_scaler__clip': False,
 'pre__commod_count_scaler__copy': True,
 'pre__commod_count_scaler__feature_range': (0, 1),
 'pre__commod_names-down_cv': CountVectorizer(),
```

```
'pre__commod_names-down_cv__analyzer': 'word',
'pre__commod_names-down_cv__binary': False,
'pre__commod_names-down_cv__decode_error': 'strict',
'pre__commod_names-down_cv__dtype': numpy.int64,
'pre__commod_names-down_cv__encoding': 'utf-8',
'pre__commod_names-down_cv__input': 'content',
'pre__commod_names-down_cv__lowercase': True,
'pre__commod_names-down_cv__max_df': 1.0,
'pre__commod_names-down_cv__max_features': None,
'pre__commod_names-down_cv__min_df': 1,
'pre__commod_names-down_cv__ngram_range': (1, 1),
'pre__commod_names-down_cv__preprocessor': None,
'pre__commod_names-down_cv__stop_words': None,
'pre__commod_names-down_cv__strip_accents': None,
'pre__commod_names-down_cv__token_pattern': '(?u)\\b\\w\\w+\\b',
'pre__commod_names-down_cv__tokenizer': None,
'pre__commod_names-down_cv__vocabulary': None,
'pre__commod_names-short_cv': CountVectorizer(),
'pre__commod_names-short_cv__analyzer': 'word',
'pre__commod_names-short_cv__binary': False,
'pre__commod_names-short_cv__decode_error': 'strict',
'pre__commod_names-short_cv__dtype': numpy.int64,
'pre__commod_names-short_cv__encoding': 'utf-8',
'pre__commod_names-short_cv__input': 'content',
'pre__commod_names-short_cv__lowercase': True,
'pre__commod_names-short_cv__max_df': 1.0,
'pre__commod_names-short_cv__max_features': None,
'pre__commod_names-short_cv__min_df': 1,
'pre__commod_names-short_cv__ngram_range': (1, 1),
'pre__commod_names-short_cv__preprocessor': None,
'pre__commod_names-short_cv__stop_words': None,
'pre__commod_names-short_cv__strip_accents': None,
'pre__commod_names-short_cv__token_pattern': '(?u)\\b\\w\\w+\\b',
'pre__commod_names-short_cv__tokenizer': None,
'pre__commod_names-short_cv__vocabulary': None,
'pre__commod_names_up_cv': CountVectorizer(),
'pre__commod_names_up_cv__analyzer': 'word',
'pre__commod_names_up_cv__binary': False,
'pre__commod_names_up_cv__decode_error': 'strict',
'pre__commod_names_up_cv__dtype': numpy.int64,
'pre__commod_names_up_cv__encoding': 'utf-8',
'pre__commod_names_up_cv__input': 'content',
'pre__commod_names_up_cv__lowercase': True,
'pre__commod_names_up_cv__max_df': 1.0,
'pre__commod_names_up_cv__max_features': None,
'pre__commod_names_up_cv__min_df': 1,
'pre__commod_names_up_cv__ngram_range': (1, 1),
```

```

'pre__commod_names_up_cv__preprocessor': None,
'pre__commod_names_up_cv__stop_words': None,
'pre__commod_names_up_cv__strip_accents': None,
'pre__commod_names_up_cv__token_pattern': '(?u)\\b\\w+\\b',
'pre__commod_names_up_cv__tokenizer': None,
'pre__commod_names_up_cv__vocabulary': None,
'pre__n_jobs': None,
'pre__remainder': 'drop',
'pre__sparse_threshold': 0.3,
'pre__summary_vectorizer': W2VTransformer(),
'pre__summary_vectorizer__alpha': 0.025,
'pre__summary_vectorizer__batch_words': 10000,
'pre__summary_vectorizer__cbow_mean': 1,
'pre__summary_vectorizer__epochs': 200,
'pre__summary_vectorizer__hashfxn': <function hash>,
'pre__summary_vectorizer__hs': 0,
'pre__summary_vectorizer__max_vocab_size': None,
'pre__summary_vectorizer__min_alpha': 0.0001,
'pre__summary_vectorizer__min_count': 2,
'pre__summary_vectorizer__negative': 5,
'pre__summary_vectorizer__null_word': 0,
'pre__summary_vectorizer__sample': 0.001,
'pre__summary_vectorizer__seed': 1,
'pre__summary_vectorizer__sg': 0,
'pre__summary_vectorizer__sorted_vocab': 1,
'pre__summary_vectorizer__trim_rule': None,
'pre__summary_vectorizer__vector_size': 20,
'pre__summary_vectorizer__window': 30,
'pre__summary_vectorizer__workers': 3,
'pre__transformer_weights': None,
'pre__transformers': [('summary_vectorizer',
    W2VTransformer(),
    'clean_summary'),
    ('wras_vectorizer', W2VTransformer(), 'clean_wras'),
    ('commod_names_up_cv', CountVectorizer(),
'commodities_up_in_price'),
    ('commod_names-down_cv', CountVectorizer(),
'commodities_down_in_price'),
    ('commod_names-short_cv', CountVectorizer(),
'commodities_in_short_supply'),
    ('commod_count_scaler',
    MinMaxScaler(),
    ['commodities_up_count',
    'commodities_down_count',
    'commodities_short_count'])],
'pre__verbose': False,
'pre__wras_vectorizer': W2VTransformer(),

```

```

'pre__wras_vectorizer__alpha': 0.025,
'pre__wras_vectorizer__batch_words': 10000,
'pre__wras_vectorizer__cbow_mean': 1,
'pre__wras_vectorizer__epochs': 200,
'pre__wras_vectorizer__hashfxn': <function hash>,
'pre__wras_vectorizer__hs': 0,
'pre__wras_vectorizer__max_vocab_size': None,
'pre__wras_vectorizer__min_alpha': 0.0001,
'pre__wras_vectorizer__min_count': 2,
'pre__wras_vectorizer__negative': 5,
'pre__wras_vectorizer__null_word': 0,
'pre__wras_vectorizer__sample': 0.001,
'pre__wras_vectorizer__seed': 1,
'pre__wras_vectorizer__sg': 0,
'pre__wras_vectorizer__sorted_vocab': 1,
'pre__wras_vectorizer__trim_rule': None,
'pre__wras_vectorizer__vector_size': 20,
'pre__wras_vectorizer__window': 30,
'pre__wras_vectorizer__workers': 3,
'sgd': SGDClassifier(alpha=1e-08, class_weight='balanced',
penalty='l1'),
'sgd__alpha': 1e-08,
'sgd__average': False,
'sgd__class_weight': 'balanced',
'sgd__early_stopping': False,
'sgd__epsilon': 0.1,
'sgd__eta0': 0.0,
'sgd__fit_intercept': True,
'sgd__l1_ratio': 0.15,
'sgd__learning_rate': 'optimal',
'sgd__loss': 'hinge',
'sgd__max_iter': 1000,
'sgd__n_iter_no_change': 5,
'sgd__n_jobs': None,
'sgd__penalty': 'l1',
'sgd__power_t': 0.5,
'sgd__random_state': None,
'sgd__shuffle': True,
'sgd__tol': 0.001,
'sgd__validation_fraction': 0.1,
'sgd__verbose': 0,
'sgd__warm_start': False,
'steps': [('pre',
ColumnTransformer(transformers=[('summary_vectorizer',
W2VTransformer(),
'clean_summary'),

```

```

W2VTransformer(),

CountVectorizer(),

CountVectorizer(),

CountVectorizer(),

MinMaxScaler(),

('wras_vectorizer',

 'clean_wras'),
('commod_names_up_cv',

 'commodities_up_in_price'),
('commod_names-down_cv',

 'commodities_down_in_price'),
('commod_names-short_cv',

 'commodities_in_short_supply'),
('commod_count_scaler',

 ['commodities_up_count',
  'commodities_down_count',
  'commodities_short_count'])),
('sgd', SGDClassifier(alpha=1e-08, class_weight='balanced',
penalty='l1'))],
'verbose': False}

```

