Tensor Flow cheat sheet

Efim Abrikosov

April 25, 2019

1 Data Prepocessing

1.1 Feature columns

`tf.feature_column`

- `tf.feature_column.numeric_column(`*colname*`)`
- `tf.feature_column.bucketized_column(`*tf.feature_coumn.numeric_column*, *boundaries_list*`)`
- `tf.feature_column.crossed_column(`*[tf.feature_coumn.categorical_column list]*, *num_hash_buckets*`)`
- `tf.feature_column.categorical_column_with_identity(`*colname*, *num_buckets*`)`
- `tf.feature_column.categorical_column_with_vocabulary_list(`*colname*, *vocabulary_list*`)`
- `tf.feature_column.categorical_column_with_hash_bucket()`*colname*, *hash_bucket_size*`)`
- `tf.feature_column.embedding_column(`*tf.feature_column.categorical_column*, *n_dim*`)`

2 Estimator API

`tf.estimator`

- `tf.estimator.Estimator(`*model_fn, out_dir*`)`

- *model_fn* is a function that returns *tf.estimator.EstimatorSpec* given a triple (*feature_cols*, *targets*, *mode*)

  Sample signature: `tf.estimator.EstimatorSpec(**args)`

  Arguments list consists of:

  1. *mode*, one of the three:
     `tf.estimator.ModeKeys.TRAIN`,
     `tf.estimator.ModeKeys.EVAL`,
     `tf.estimator.ModeKeys.PREDICT`
  2. *predictions_dict*
  3. *loss*
  4. *train_op*
  5. *eval_metrics_ops*
  6. *export_outputs*

- *model_dir* is the name of output directory (string)

- `tf.estimator.train_and_evaluate(estimator, train_spec, eval_spec)`

  - *estimator* is an instance of *tf.estimator*

    Sample singature: `tf.estimator.[model_name](model_dir, feature_columns, model_params)`
  - *train_spec* is an instance of *tf.estimator.TrainSpec*

    Sample signature: `tf.estimator.TrainSpec(input_fn, max_steps)`
  - *eval_spec* is an instance of *tf.estimator.EvalSpec*

    Sample signature: `tf.estimator.EvalSpec(input_fn,steps, start_delay_secs, throttle_secs)`

- Adding performance metrics

  Sample signature: `estimator = tf.contrib.estimator.add_metrics(estimator,rmse)`