

The validation scheme is based on [pschmitt/mw-with-gn](https://github.com/pschmitt/mw-with-gn), and cleaned data is from [data/without-gn](https://github.com/pschmitt/mw-without-gn) and Kalman filter is from <https://www.kaggle.com/bernatmag02/kaggle-model-150kmn-kalman-filter> and the added feature is from [wavenet/with-gn-new-feature](https://github.com/pschmitt/mw-with-gn-new-feature). I also used ragnar's data in this version [clean_kalman](https://github.com/pschmitt/mw-with-gn). The Wavenet is based on <https://github.com/ebelloperem/beras-dec>, <https://github.com/pschmitt/wavenet> and <https://github.com/pschmitt/wavenet> and also <https://www.kaggle.com/wimwim/wavenet-150km>. If any reference is not mentioned it was not intentional, please add them in comments.

Previous versions were mainly based on <https://www.kaggle.com/wimwim/wavenet-150km>

```

Cell, SimpleNN, TimeDistributed, RM,
        Add, AveragePooling1D, Multiply, GRU, GRUCell, LSTMCell, SimpleRNN,
        RepeatVector, Conv1D, MaxPooling1D, Concatenate, GlobalAveragePool
ing1D, UpSampling1D)
from tensorflow.keras.callbacks import ModelCheckpoint, EarlyStopping, Callback, ReduceLROnPlateau, LearningRateScheduler
from tensorflow.keras.losses import binary_crossentropy, categorical_crossentropy, mean_squared_error
from tensorflow.keras.optimizers import Adam, RMSprop, SGD
from tensorflow.keras.utils import Sequence, to_categorical
from sklearn.model_selection import KFold, GroupKFold
from tensorflow.keras import backend as K
import tensorflow as tf
from typing import List, NoReturn, Union, Tuple, Optional, Text, Generic, Callable, Dict
from sklearn.metrics import f1_score, cohen_kappa_score, mean_squared_error
from logging import getLogger, Formatter, StreamHandler, FileHandler, INFO
from sklearn.metrics import selection import f1_score, GroupKFold
from tqdm import tqdm, tqdm_notebook as tqdm
from contextlib import contextmanager
from joblib import Parallel, delayed
from IPython.display import display
from sklearn import preprocessing
import tensorflow_addons as tfa
import scipy.stats as stats
import random as rn
import pandas as pd
import numpy as np
import scipy as sp
import itertools
import warnings
import time
import pywt
import os
import gc

warnings.simplefilter('ignore')
warnings.filterwarnings('ignore')
pd.set_option('display.max_columns', 1000)
pd.set_option('display.max_rows', 500)
matplotlib inline

EPOCHS=120
NNBATCHES=20
BATCHSIZE = 4000
SEED = 321
SELECT = True
SPLITS = 5
LR = 0.001
fe_config = {
    (True, 4000),
}

def init_logger():
    handler = StreamHandler()
    handler.setLevel(INFO)
    handler.setFormatter(Formatter(LOGFORMAT))
    fh_handler = FileHandler('{}.log'.format(MODELNAME))
    fh_handler.setFormatter(Formatter(LOGFORMAT))
    logger.setLevel(INFO)
    logger.addHandler(handler)
    logger.addHandler(fh_handler)

@contextmanager
def timer(name : Text):
    t0 = time.time()
    yield
    logger.info(f'[{name}] done in {time.time() - t0:.0f} s')

COMPETITION = 'ION-Switching'
logger = getLogger(COMPETITION)
LOGFORMAT = '%(asctime)s : %(levelname)s : %(message)s'
MODELNAME = 'WaveNet'

def seed_everything(seed : int) -> NoReturn :
    rn.seed(seed)
    np.random.seed(seed)
    os.environ['PYTHONHASHSEED'] = str(seed)
    tf.random.set_seed(seed)

seed_everything(SEED)

def read_data(base : os.path.abspath) -> Tuple[pd.DataFrame, pd.DataFrame, pd.DataFrame]:
    train = pd.read_csv('%kaggle/input/clean-kalman/train_clean_kalman.csv', dtype={'time': np.float32,
'signal': np.float32}, 'open_channels'> np.int32))
    test = pd.read_csv('%kaggle/input/clean-kalman/test_clean_kalman.csv', dtype={'time': np.float32,
'signal': np.float32})
    sub = pd.read_csv('%kaggle/input/liverpool-ion-switching/sample_submission.csv', dtype={'time': np
.float32})
    return train, test, sub

def batching(df : pd.DataFrame,
            batch_size : int) -> pd.DataFrame :
    df['group'] = df.groupby(df.index//batch_size, sort=False)['signal'].agg(['ngroup']).values
    df['group'] = df['group'].astype(np.uint6)
    return df

def reduce_mem_usage(df : pd.DataFrame,
                    verbose: bool = True) -> pd.DataFrame:
    numerics = ['int16', 'int32', 'int64', 'float16', 'float32', 'float64']
    start_mem = df.memory_usage().sum() / 1024**2
    for col in df.columns:
        col_type = df[col].dtypes
        if col_type in numerics:
            c_min = df[col].min()
            c_max = df[col].max()
            if str(col_type)[:3] == 'int':
                if c_min > np.iinfo(np.int32).min
                    and c_max < np.iinfo(np.int32).max):
                    df[col] = df[col].astype(np.int32)
                elif c_min > np.iinfo(np.int64).min
                    and c_max < np.iinfo(np.int64).max):
                    df[col] = df[col].astype(np.int64)
            else:
                if c_min > np.finfo(np.float16).min
                    and c_max < np.finfo(np.float16).max):
                    df[col] = df[col].astype(np.float16)
                elif c_min > np.finfo(np.float32).min
                    and c_max < np.finfo(np.float32).max):
                    df[col] = df[col].astype(np.float32)
                else:
                    df[col] = df[col].astype(np.float64)
    end_mem = df.memory_usage().sum() / 1024**2
    reduction = (start_mem - end_mem) / start_mem
    msg = f'Mem. usage decreased to {end_mem:.5f} MB ({reduction * 100:.1f} % reduction)'
    if verbose:
        print(msg)
    return df

def lag_with_pct_change(df : pd.DataFrame,
                        shift_sizes : Optional[List]= [1, 2],
                        add_pct_change : Optional[bool]=False,
                        add_pct_change_lag : Optional[bool]=False) -> pd.DataFrame:
    for shift_size in shift_sizes:
        df[f'lag_{shift_size}'] = df.groupby('group').shift(shift_size)
        if add_pct_change:
            df[f'pct_change_{shift_size}'] = df[f'lag_{shift_size}'].pct_change()
        if add_pct_change_lag:
            df[f'pct_change_lag_{shift_size}'] = df[f'pct_change_{shift_size}'].lag(shift_size)

```

```
a(0)
df['signal_shift_neg_'+str(shift_size)] = df.groupby('group')['signal'].shift(-1*shift_size).fillna(0)
```


1241:run_everything(flg)

2020-04-04 15:47:22,700 INFO Reading Data Started ...

2020-04-04 15:47:22,700 INFO Reading and Normalizing Data Completed ...

2020-04-04 15:47:22,703 INFO [Reading Data] done in 3 s

2020-04-04 15:47:25,540 INFO Feature Engineering Completed

2020-04-04 15:47:35,542 INFO [Creating Features] done in 13 s

2020-04-04 15:47:35,543 INFO Training Wavenet model with 5 folds of GroupKFold Started ...

FlMacro: 0.93717

FlMacro: 0.92537

FlMacro: 0.93041

FlMacro: 0.93231

FlMacro: 0.93446

FlMacro: 0.93602

FlMacro: 0.93300

FlMacro: 0.93566

FlMacro: 0.93413

FlMacro: 0.93416

FlMacro: 0.93263

FlMacro: 0.93659

FlMacro: 0.93185

FlMacro: 0.93369

FlMacro: 0.93742

FlMacro: 0.91941

FlMacro: 0.93689

FlMacro: 0.93720

FlMacro: 0.93569

FlMacro: 0.93411

FlMacro: 0.93578

FlMacro: 0.93499

FlMacro: 0.93756

FlMacro: 0.93739

FlMacro: 0.93156

FlMacro: 0.93788

FlMacro: 0.92603

FlMacro: 0.93644

FlMacro: 0.93475

FlMacro: 0.93781

FlMacro: 0.93754

FlMacro: 0.93808

FlMacro: 0.93602

FlMacro: 0.93753

FlMacro: 0.93798

FlMacro: 0.93832

FlMacro: 0.93265

FlMacro: 0.93829

FlMacro: 0.93776

FlMacro: 0.93832

FlMacro: 0.93844

FlMacro: 0.93849

FlMacro: 0.93863

FlMacro: 0.93850

FlMacro: 0.93833

FlMacro: 0.93836

FlMacro: 0.93817

FlMacro: 0.93852

FlMacro: 0.93862

FlMacro: 0.93668

FlMacro: 0.93867

FlMacro: 0.93864

FlMacro: 0.93866

FlMacro: 0.93867

FlMacro: 0.93884

FlMacro: 0.93870

FlMacro: 0.93844

FlMacro: 0.93810

FlMacro: 0.93882

FlMacro: 0.93865

FlMacro: 0.93836

FlMacro: 0.93862

FlMacro: 0.93873

FlMacro: 0.93880

FlMacro: 0.93875

FlMacro: 0.93872

FlMacro: 0.93871

FlMacro: 0.93868

FlMacro: 0.93871

FlMacro: 0.93860

FlMacro: 0.93891

FlMacro: 0.93882

FlMacro: 0.93882

FlMacro: 0.93881

FlMacro: 0.93881

FlMacro: 0.93874

FlMacro: 0.93874

FlMacro: 0.93887

FlMacro: 0.93865

FlMacro: 0.93885

FlMacro: 0.93870

FlMacro: 0.93858

FlMacro: 0.93858

FlMacro: 0.93875

FlMacro: 0.93889

FlMacro: 0.93885

FlMacro: 0.93870

FlMacro: 0.93881

FlMacro: 0.93891

FlMacro: 0.93872

FlMacro: 0.93888

FlMacro: 0.93877

FlMacro: 0.93886

FlMacro: 0.93880

FlMacro: 0.93856

FlMacro: 0.93881

FlMacro: 0.93886

FlMacro: 0.93876

FlMacro: 0.93894

FlMacro: 0.93877

FlMacro: 0.93887

FlMacro: 0.93892

FlMacro: 0.93878

FlMacro: 0.93891

FlMacro: 0.93883

FlMacro: 0.93883

FlMacro: 0.93878

FlMacro: 0.93880

FlMacro: 0.93878

FlMacro: 0.93879

FlMacro: 0.93882

FlMacro: 0.93887

FlMacro: 0.93862

FlMacro: 0.93881

FlMacro: 0.93881

2020-04-04 16:20:26,858 INFO Training fold 1 completed. macro f1 score : 0.93881

FlMacro: 0.89838

FlMacro: 0.92962

FlMacro: 0.93724

FlMacro: 0.93647

FlMacro: 0.93746

FlMacro: 0.93810

FlMacro: 0.93554

FlMacro: 0.93801

FlMacro: 0.93672

FlMacro: 0.92553

FlMacro: 0.93760

FlMacro: 0.93856

FlMacro: 0.93743

FlMacro: 0.93883

FlMacro: 0.93856

FlMacro: 0.84923

FlMacro: 0.84762

FlMacro: 0.84855

FlMacro: 0.84835

FlMacro: 0.84913

FlMacro: 0.84753

FlMacro: 0.84930

FlMacro: 0.84924

FlMacro: 0.84929

FlMacro: 0.84869

FlMacro: 0.84949

FlMacro: 0.84906

FlMacro: 0.84969

FlMacro: 0.84943

FlMacro: 0.92684

FlMacro: 0.93836

FlMacro: 0.93903

FlMacro: 0.93870

FlMacro: 0.93890

FlMacro: 0.93940

FlMacro: 0.93861

FlMacro: 0.93914

FlMacro: 0.93817

FlMacro: 0.93934

FlMacro: 0.93855

FlMacro: 0.93983

FlMacro: 0.93973

FlMacro: 0.93976

FlMacro: 0.93971

FlMacro: 0.93978

FlMacro: 0.93991

FlMacro: 0.93974

FlMacro: 0.93927

FlMacro: 0.93957

FlMacro: 0.93976

FlMacro: 0.93982

FlMacro: 0.93979

FlMacro: 0.93989

FlMacro: 0.93962

FlMacro: 0.93961

FlMacro: 0.93969

FlMacro: 0.93985

FlMacro: 0.93966

FlMacro: 0.93974

FlMacro: 0.93977

FlMacro: 0.93988

FlMacro: 0.93992

FlMacro: 0.93990

FlMacro: 0.94002

FlMacro: 0.93996

FlMacro: 0.93972

FlMacro: 0.93972

FlMacro: 0.93979

FlMacro: 0.93972

FlMacro: 0.93984

FlMacro: 0.93984

FlMacro: 0.93959

FlMacro: 0.93983

FlMacro: 0.93963

FlMacro: 0.93970

FlMacro: 0.93972

FlMacro: 0.93983

FlMacro: 0.93966

FlMacro: 0.93981

FlMacro: 0.93984

FlMacro: 0.93979

FlMacro: 0.93975

FlMacro: 0.93981

FlMacro: 0.93988

FlMacro: 0.93982

FlMacro: 0.93996

FlMacro: 0.93987

FlMacro: 0.93997

FlMacro: 0.93974

FlMacro: 0.93990

FlMacro: 0.94001

FlMacro: 0.93992

FlMacro: 0.94003

FlMacro: 0.93994

FlMacro: 0.93990

FlMacro: 0.93978

FlMacro: 0.93984

FlMacro: 0.93984

FlMacro: 0.93986

FlMacro: 0.93990

FlMacro: 0.93989

2020-04-04 17:11:36,492 INFO Training fold 2 completed. macro f1 score : 0.93989

FlMacro: 0.78224

FlMacro: 0.90813

FlMacro: 0.92552

FlMacro: 0.92897

FlMacro: 0.93046

FlMacro: 0.93151

FlMacro: 0.93454

FlMacro: 0.89066

FlMacro: 0.93234

FlMacro: 0.93297

FlMacro: 0.93624

FlMacro: 0.93550

FlMacro: 0.93595

FlMacro: 0.93888

FlMacro: 0.93300

FlMacro: 0.92601

FlMacro: 0.93593

FlMacro: 0.93376

FlMacro: 0.93604

FlMacro: 0.93785

FlMacro: 0.93516

FlMacro: 0.93473

FlMacro: 0.93585

FlMacro: 0.93618

FlMacro: 0.93675

FlMacro: 0.93637

FlMacro: 0.93664

FlMacro: 0.93406

FlMacro: 0.93569

FlMacro: 0.93606

FlMacro: 0.93667

FlMacro: 0.93656

FlMacro: 0.90113

FlMacro: 0.93540

FlMacro: 0.93459

FlMacro: 0.93609

FlMacro: 0.93624

FlMacro: 0.93668

FlMacro: 0.93625

FlMacro: 0.93716

FlMacro: 0.93712

FlMacro: 0.93737

FlMacro: 0.93718

FlMacro: 0.93719

FlMacro: 0.93723

FlMacro: 0.93736

FlMacro: 0.93719

FlMacro: 0.93732

FlMacro: 0.93730

FlMacro: 0.93728

FlMacro: 0.93731

FlMacro: 0.93738

FlMacro: 0.93730

FlMacro: 0.93740

FlMacro: 0.93741

FlMacro: 0.93728

FlMacro: 0.93727

FlMacro: 0.93728

FlMacro: 0.93738

FlMacro: 0.93743

FlMacro: 0.93739

FlMacro: 0.93745

FlMacro: 0.93754

FlMacro: 0.93742

FlMacro: 0.93730

FlMacro: 0.93740

FlMacro: 0.93741

FlMacro: 0.93742

FlMacro: 0.93731

FlMacro: 0.93742

FlMacro: 0.93741

FlMacro: 0.93754

FlMacro: 0.93751

FlMacro: 0.93755

FlMacro: 0.93745

FlMacro: 0.93746

FlMacro: 0.93747

FlMacro: 0.93743

FlMacro: 0.93746

FlMacro: 0.93749

FlMacro: 0.93749

FlMacro: 0.93754

FlMacro: 0.93740

FlMacro: 0.93760

FlMacro: 0.93754

FlMacro: 0.93740

FlMacro: 0.93762

FlMacro: 0.93764

FlMacro: 0.93740

FlMacro: 0.93746

FlMacro: 0.93736

FlMacro: 0.93758

FlMacro: 0.93762

FlMacro: 0.93763

FlMacro: 0.93752

FlMacro: 0.93760

FlMacro: 0.93769

FlMacro: 0.93764

FlMacro: 0.93751

FlMacro: 0.93750

FlMacro: 0.93775

FlMacro: 0.93766

FlMacro: 0.93763

FlMacro: 0.93768

FlMacro: 0.93768

FlMacro: 0.93763

FlMacro: 0.93759

FlMacro: 0.93760

FlMacro: 0.93760

FlMacro: 0.93763

2020-04-04 17:11:36,492 INFO Training fold 3 completed. macro f1 score : 0.93763

FlMacro: 0.78343

FlMacro: 0.85404

FlMacro: 0.92447

FlMacro: 0.92819

FlMacro: 0.92851

FlMacro: 0.93275

FlMacro: 0.93361

FlMacro: 0.93877

FlMacro: 0.93831

FlMacro: 0.93567

FlMacro: 0.93247

FlMacro: 0.93047

FlMacro: 0.93024

FlMacro: 0.93381

FlMacro: 0.93723

FlMacro: 0.93387

FlMacro: 0.93366

FlMacro: 0.93742

FlMacro: 0.93742

FlMacro: 0.93625

FlMacro: 0.93802

FlMacro: 0.93766

FlMacro: 0.93791

FlMacro: 0.93802

FlMacro: 0.93803

FlMacro: 0.93798

FlMacro: 0.93792

FlMacro: 0.93775

FlMacro: 0.93773

FlMacro: 0.93756

FlMacro: 0.93791

FlMacro: 0.93794

FlMacro: 0.93792

FlMacro: 0.93799

FlMacro: 0.93795

FlMacro: 0.93820

FlMacro: 0.93771

FlMacro: 0.93802

FlMacro: 0.93801

FlMacro: 0.93789

FlMacro: 0.93807

FlMacro: 0.93804

FlMacro: 0.93811

FlMacro: 0.93805

FlMacro: 0.93824

FlMacro: 0.93788

FlMacro: 0.93813

FlMacro: 0.93802

FlMacro: 0.93809

FlMacro: 0.93798

FlMacro: 0.93791

FlMacro: 0.93796

FlMacro: 0.93787

FlMacro: 0.93804

FlMacro: 0.93807

FlMacro: 0.93817

FlMacro: 0.93812

FlMacro: 0.93789

FlMacro: 0.93812

FlMacro: 0.93749

FlMacro: 0.93764

FlMacro: 0.93740

FlMacro: 0.93814

FlMacro: 0.93810

FlMacro: 0.93812

FlMacro: 0.93795

FlMacro: 0.93826

FlMacro: 0.93793

FlMacro: 0.93810

FlMacro: 0.93824

FlMacro: 0.93814

FlMacro: 0.93813

FlMacro: 0.93800

FlMacro: 0.93819

FlMacro: 0.93797

FlMacro: 0.93797

FlMacro: 0.93793

FlMacro: 0.93791

FlMacro: 0.93816

FlMacro: 0.93808

FlMacro: 0.93818

FlMacro: 0.93821

FlMacro: 0.93814

FlMacro: 0.93805

FlMacro: 0.93810

FlMacro: 0.93814

FlMacro: 0.93820

FlMacro: 0.93813

FlMacro: 0.93819

FlMacro: 0.93811

FlMacro: 0.93808

FlMacro: 0.93821

FlMacro: 0.93804

FlMacro: 0.93807

FlMacro: 0.93823

FlMacro: 0.93809

2020-04-04 17:15:01,226 INFO Training fold 4 completed. macro f1 score : 0.93809

FlMacro: 0.72163

FlMacro: 0.87476

FlMacro: 0.92538

FlMacro: 0.93057

FlMacro: 0.93177

FlMacro: 0.92589

FlMacro: 0.93438

FlMacro: 0.92904

FlMacro: 0.93569

FlMacro: 0.93543

FlMacro: 0.93618

FlMacro: 0.93543

FlMacro: 0.93587

FlMacro: 0.93642

FlMacro: 0.93595

FlMacro: 0.93528

FlMacro: 0.93519

FlMacro: 0.93704

FlMacro: 0.93698

FlMacro: 0.93725

FlMacro: 0.92647

FlMacro: 0.93681

FlMacro: 0.93687

FlMacro: 0.93725

FlMacro: 0.93447

FlMacro: 0.93565

FlMacro: 0.93742

FlMacro: 0.93704

FlMacro: 0.93720

FlMacro: 0.93678

FlMacro: 0.93721

FlMacro: 0.93548

FlMacro: 0.93535

FlMacro: 0.93737

FlMacro: 0.93711

FlMacro: 0.93646

FlMacro: 0.92626

FlMacro: 0.93773

FlMacro: 0.93776

FlMacro: 0.93666

FlMacro: 0.93808

FlMacro: 0.93734

FlMacro: 0.93800

FlMacro: 0.93784

FlMacro: 0.93819

FlMacro: 0.93815

FlMacro: 0.93810

FlMacro: 0.93781

FlMacro: 0.93816

FlMacro: 0.93806

FlMacro: 0.93821

FlMacro: 0.93830

FlMacro: 0.93812

FlMacro: 0.93820

FlMacro: 0.93799

FlMacro: 0.93779

FlMacro: 0.93815

FlMacro: 0.93816

FlMacro: 0.93818

FlMacro: 0.93760

FlMacro: 0.93830

FlMacro: 0.93811

FlMacro: 0.93826

FlMacro: 0.93822

FlMacro: 0.93828

FlMacro: 0.93821

FlMacro: 0.93826

FlMacro: 0.93828

FlMacro: 0.93816

FlMacro: 0.93787

FlMacro: 0.93824

FlMacro: 0.93809

FlMacro: 0.93825

FlMacro: 0.93822

FlMacro: 0.93822

FlMacro: 0.93837

FlMacro: 0.93829

FlMacro: 0.93831

FlMacro: 0.93825

FlMacro: 0.93819

FlMacro: 0.93825

FlMacro: 0.93827

FlMacro: 0.93814

FlMacro: 0.93834

FlMacro: 0.93808

FlMacro: 0.93831

FlMacro: 0.93828

FlMacro: 0.93829

FlMacro: 0.93834

FlMacro: 0.93827

FlMacro: 0.93828

FlMacro: 0.93836

FlMacro: 0.93797

FlMacro: 0.93815

FlMacro: 0.93835

FlMacro: 0.93838

FlMacro: 0.93828

FlMacro: 0.93833

FlMacro: 0.93836

FlMacro: 0.93833

FlMacro: 0.93833

FlMacro: 0.93832

FlMacro: 0.93828

FlMacro: 0.93832

FlMacro: 0.93830

FlMacro: 0.93829

FlMacro: 0.93831

FlMacro: 0.93832

FlMacro: 0.93836

FlMacro: 0.93837

FlMacro: 0.93824

FlMacro: 0.93841

2020-04-04 17:48:22,611 INFO Training fold 5 completed. macro f1 score : 0.93841

2020-04-04 17:48:25,248 INFO Training completed. oof macro f1 score : 0.93866

time	open_channels
0 500.000092	0
1 500.000214	0
2 500.000305	0
3 500.000097	0
4 500.000488	0

2020-04-04 17:48:37,654 INFO Training completed ...

2020-04-04 17:48:37,656 INFO [Running Wavenet model] done in 7262 s

In [] :