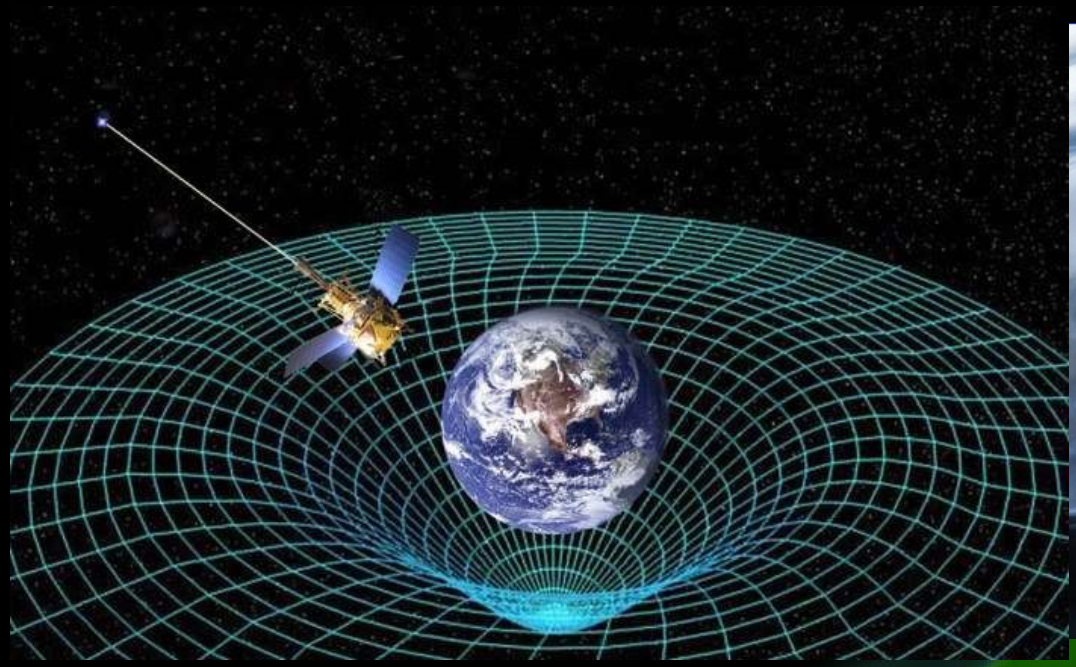
Application of machine learning in gravitational waves survey

Filip Morawski
NCAC PAS, VIRGO

GPU DAY
22.06.2018, Budapest, Hungary

Gravitational waves

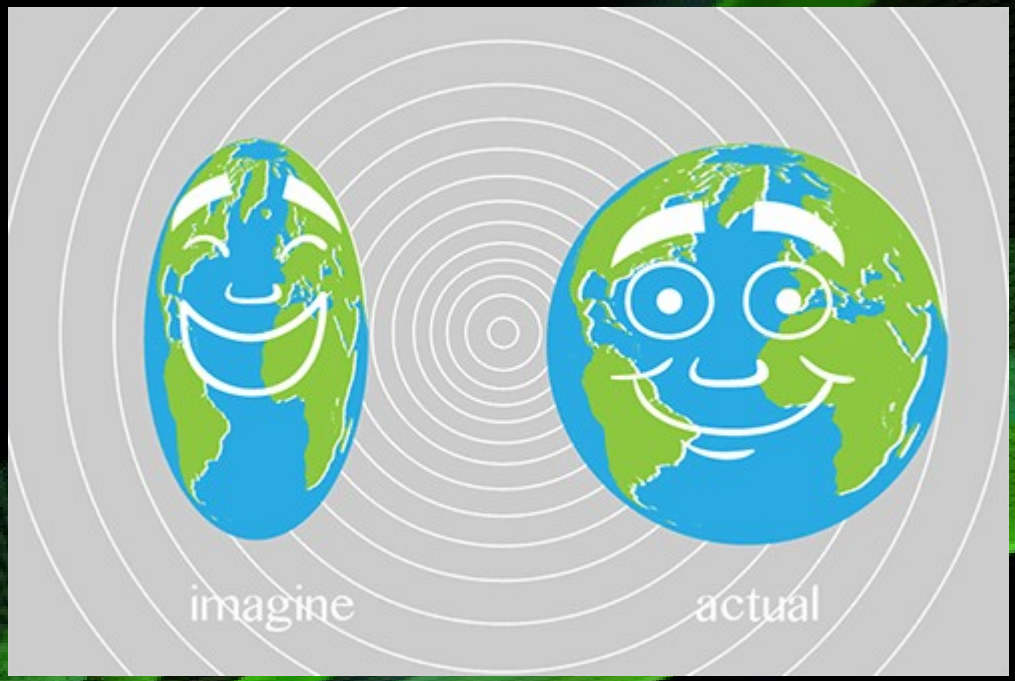


Source: LIGO Caltech

Something special happens when two bodies orbit each other – their movement cause ripples in a spacetime - like ripples in a pond after tossing a stone.

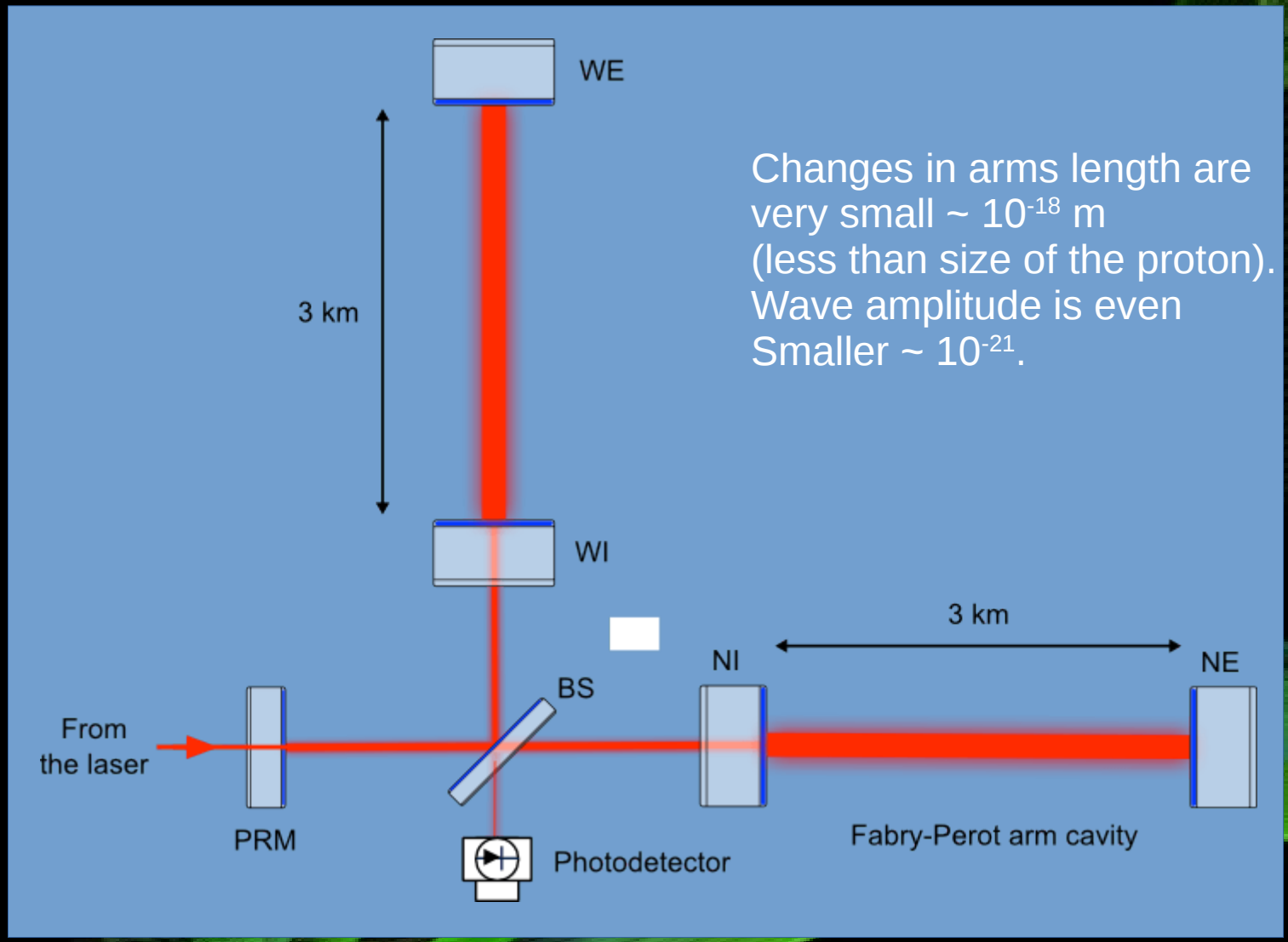
Propagation of GW

They stretch and squeeze anything in their path but in a such a small scale that it is nearly impossible to detect them.



Source: China Features

Detection principle: Laser interferometry



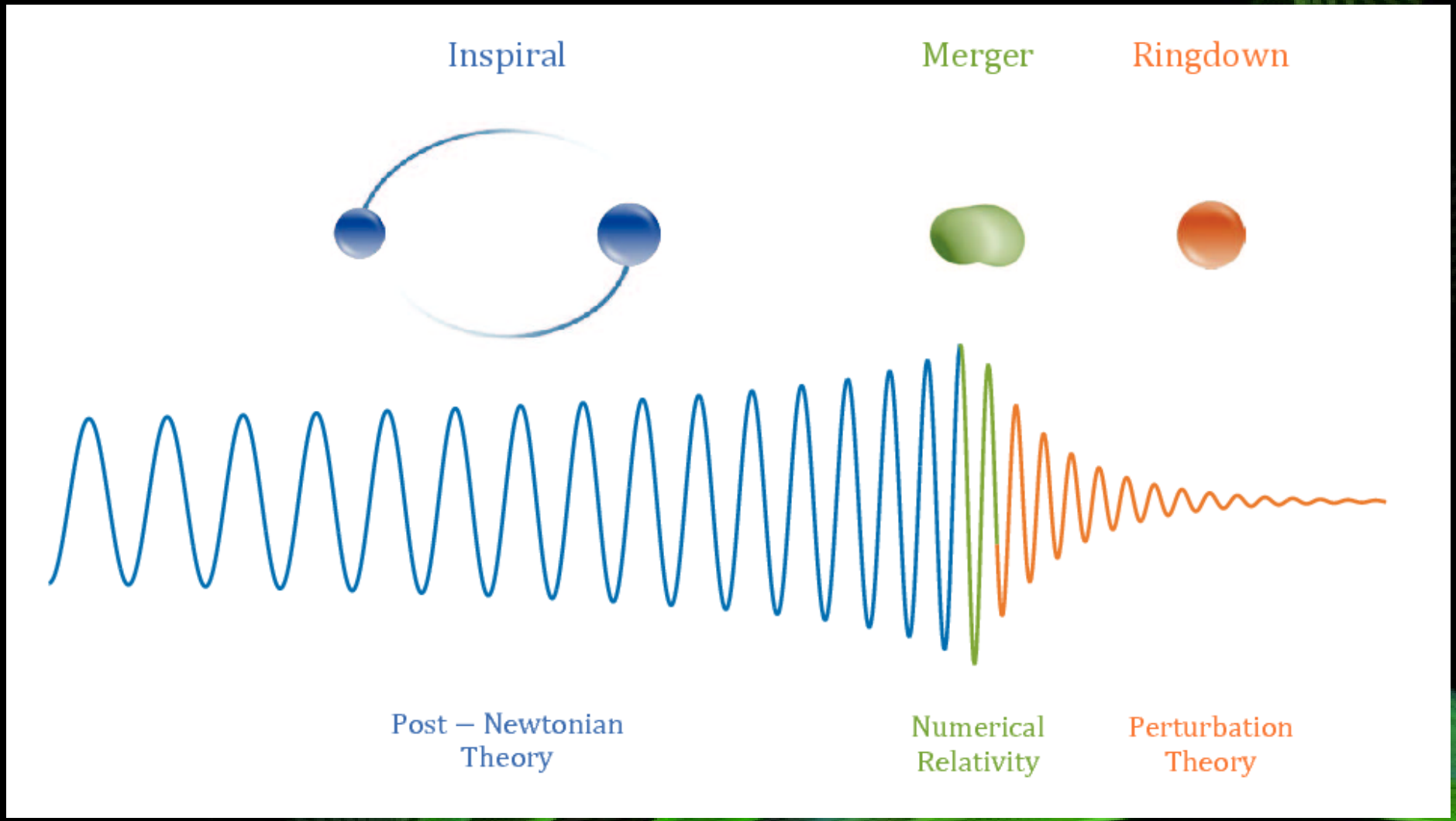
Changes in arms length are very small ~ 10^{-18} m (less than size of the proton). Wave amplitude is even Smaller ~ 10^{-21}.

Source: virgo-gw.eu

Types of signals

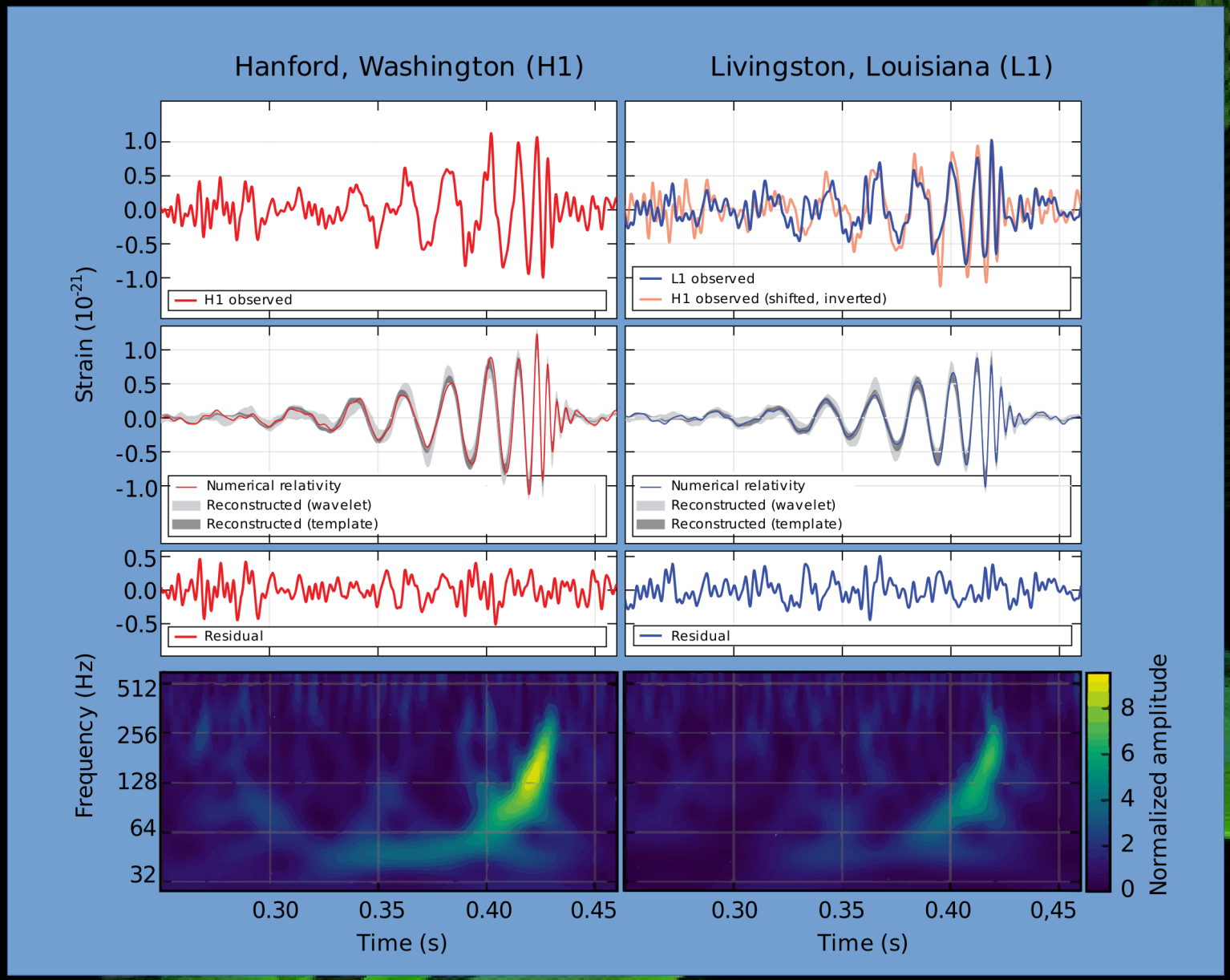
- One time cataclysmic events – mergers of binary systems (BHBH, NSNS, BHNS)
- Periodic phenomena – rotating non-axisymmetric neutron stars, wide binary systems (power of magnitudes weaker)

Evolution of binary system

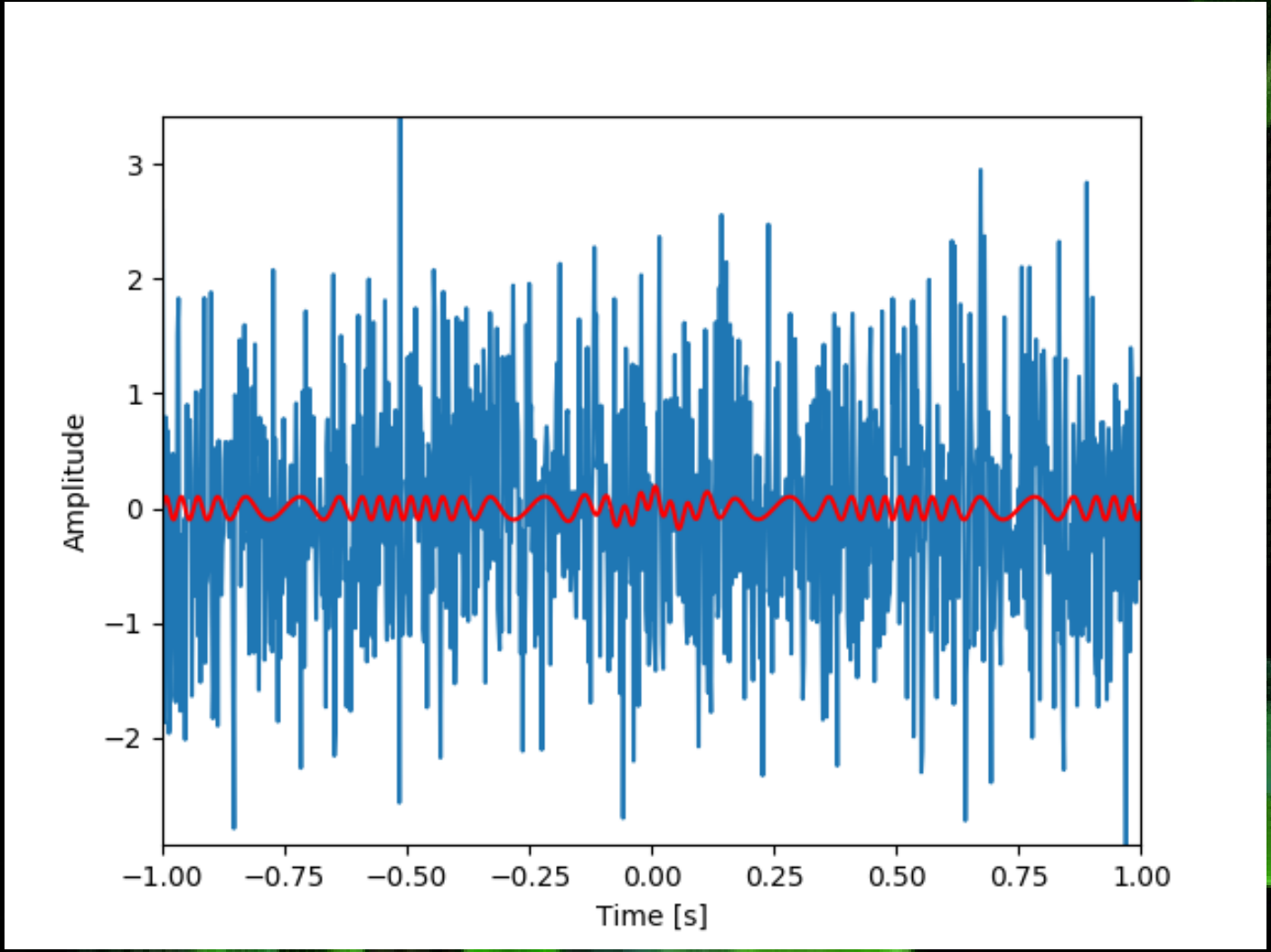


Source: Obtaining gravitational waves from inspiral binary systems using LIGO data; Antelis, J. et al. 2016

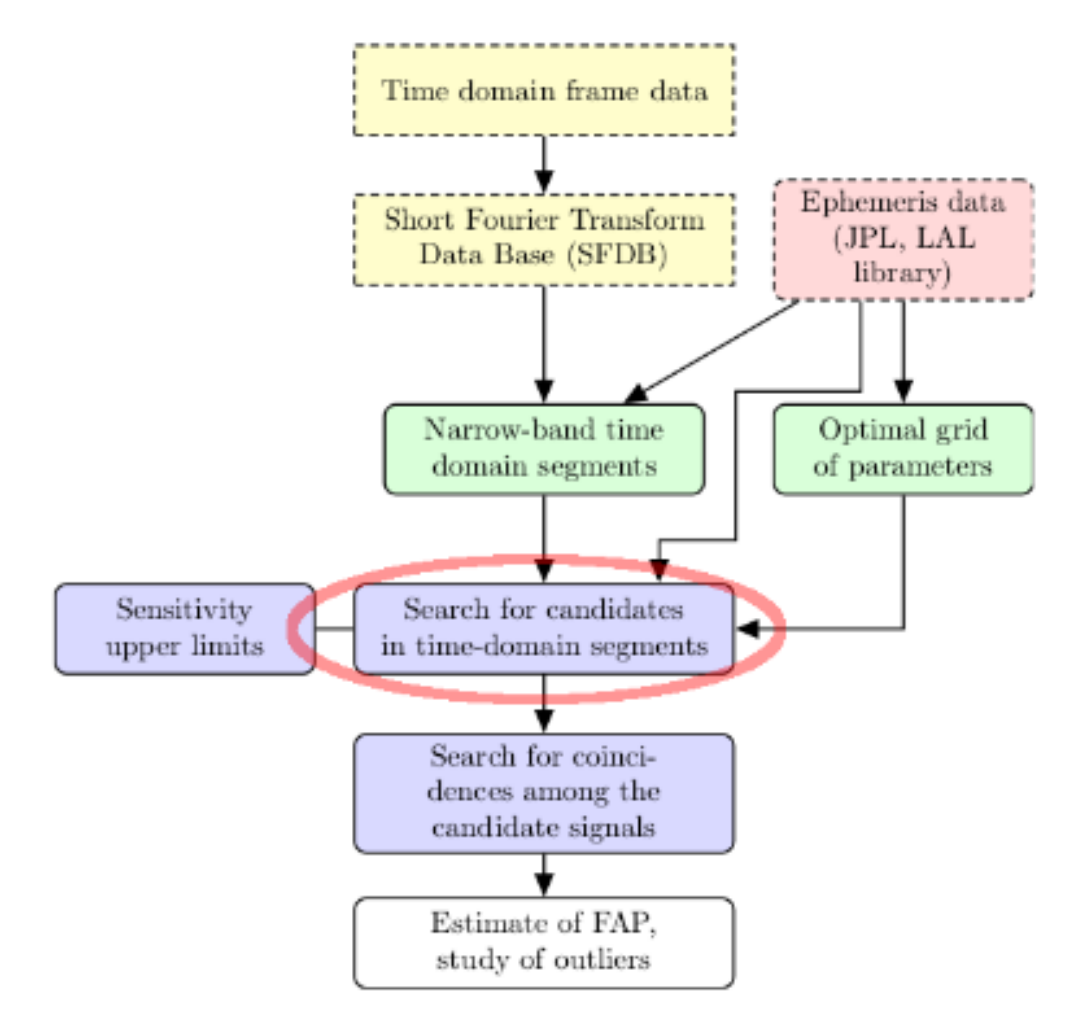
Current discoveries



Challenge



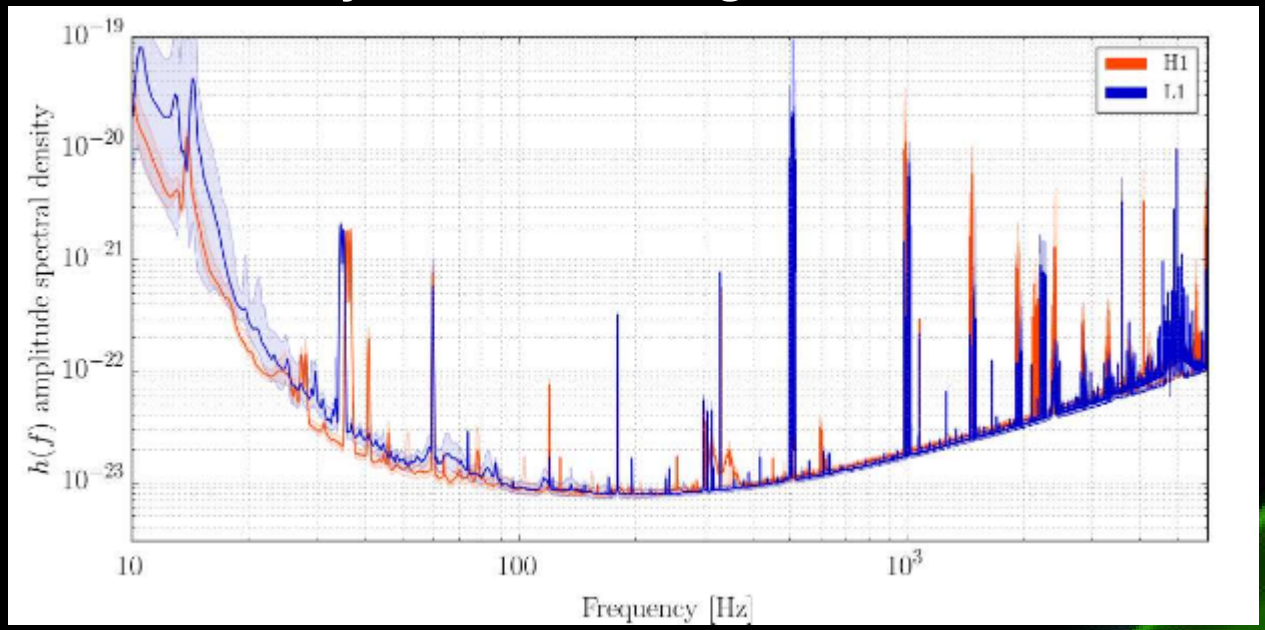
F-statistics



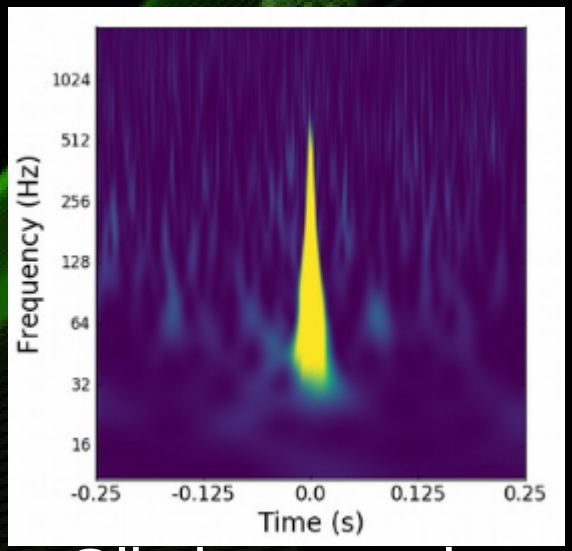
Source:github.com/mbejger/polgraw-allsky.git

Why Machine Learning?

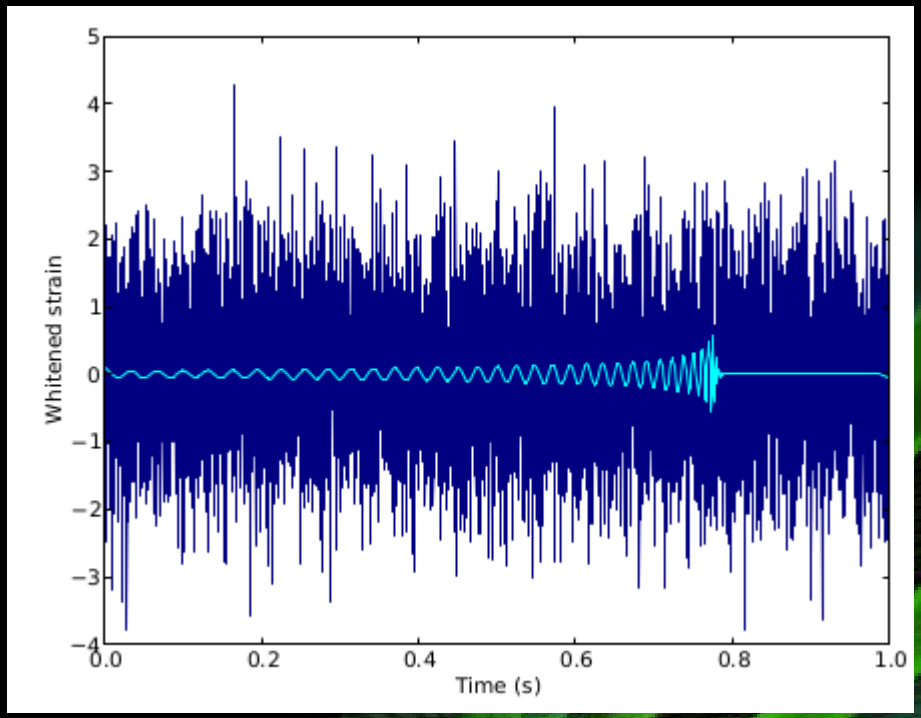
Sensitivity curve of Virgo



Source: virgo-gw.eu



Glitch – transient noise event



Time series signal of Binary Black Hole

Source: Gabbard, H. et al., 2017

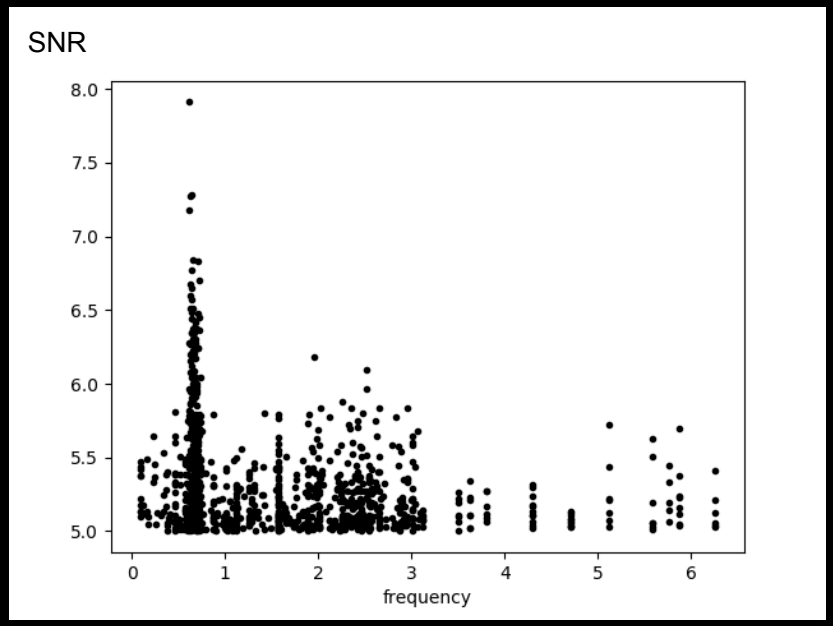
Motivation

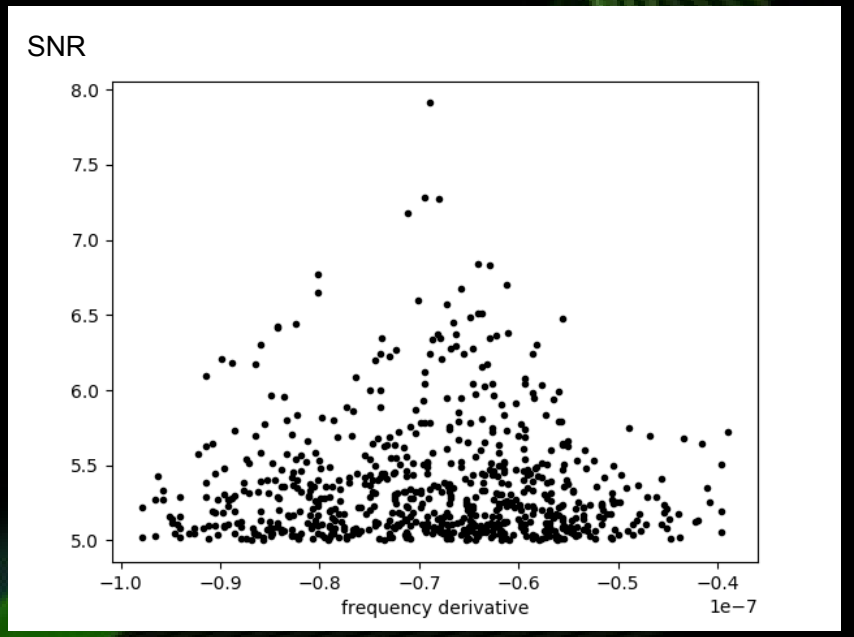
- Promising tool in classification of complex patterns and weak signals
- It may be essential for the classification of periodic signals!
- Alternative for Fourier transform based methods
- Noise reduction of unknown sources

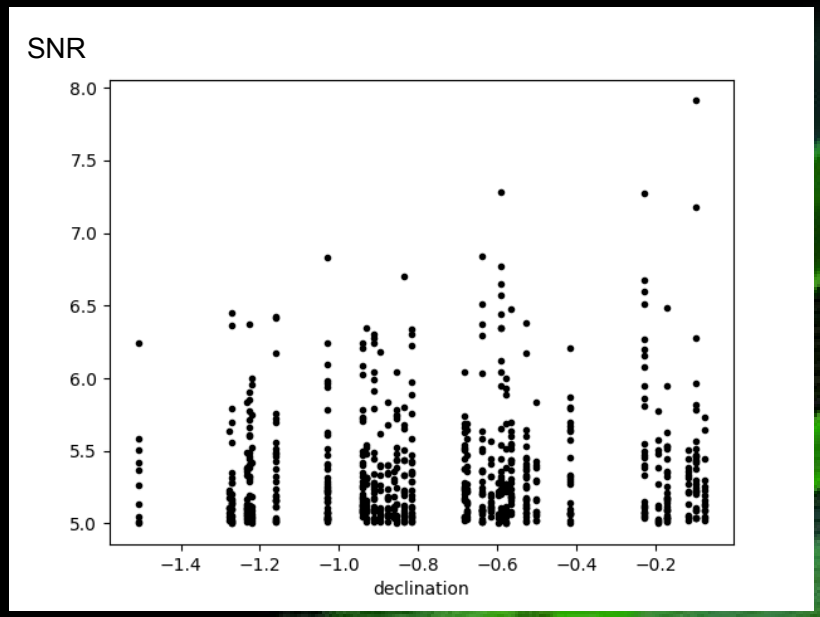
F-statistics candidates classification

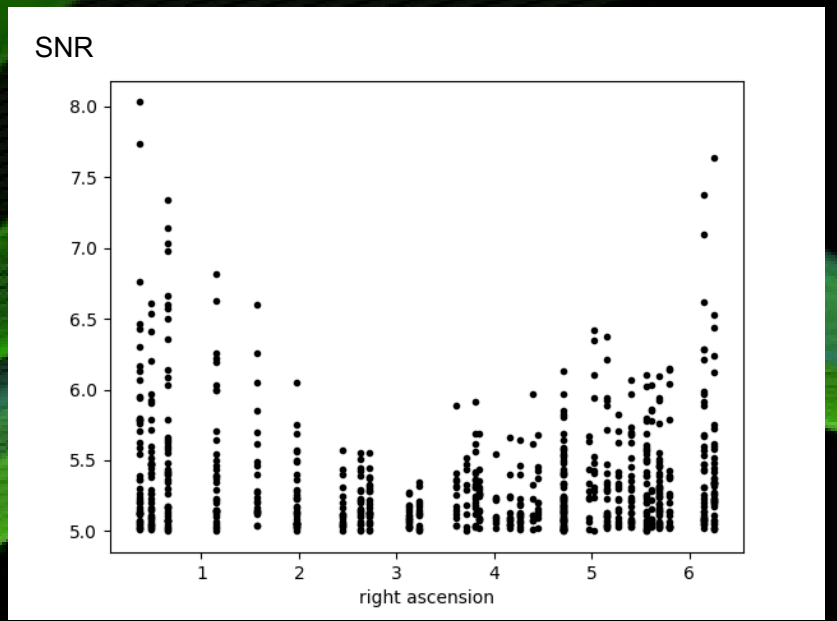
- Convolutional neural networks – search for patterns in 2D data (images)
- Data sets containing various SNRs and frequency of the signal (various number of points)
- Our own definition of SNR – integrated over whole signal; increases both with width of time window and amplitude of the signal

Generated data

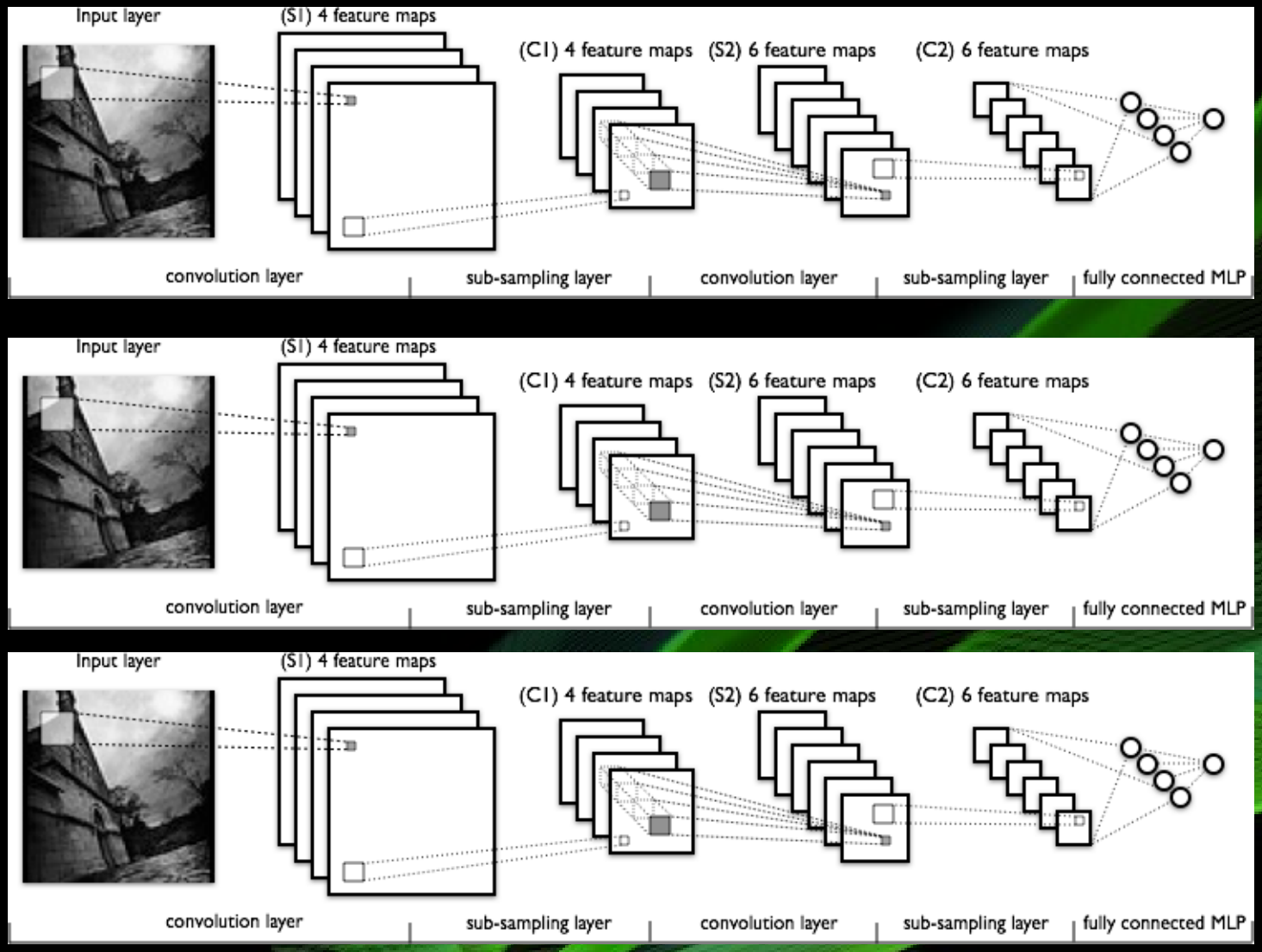






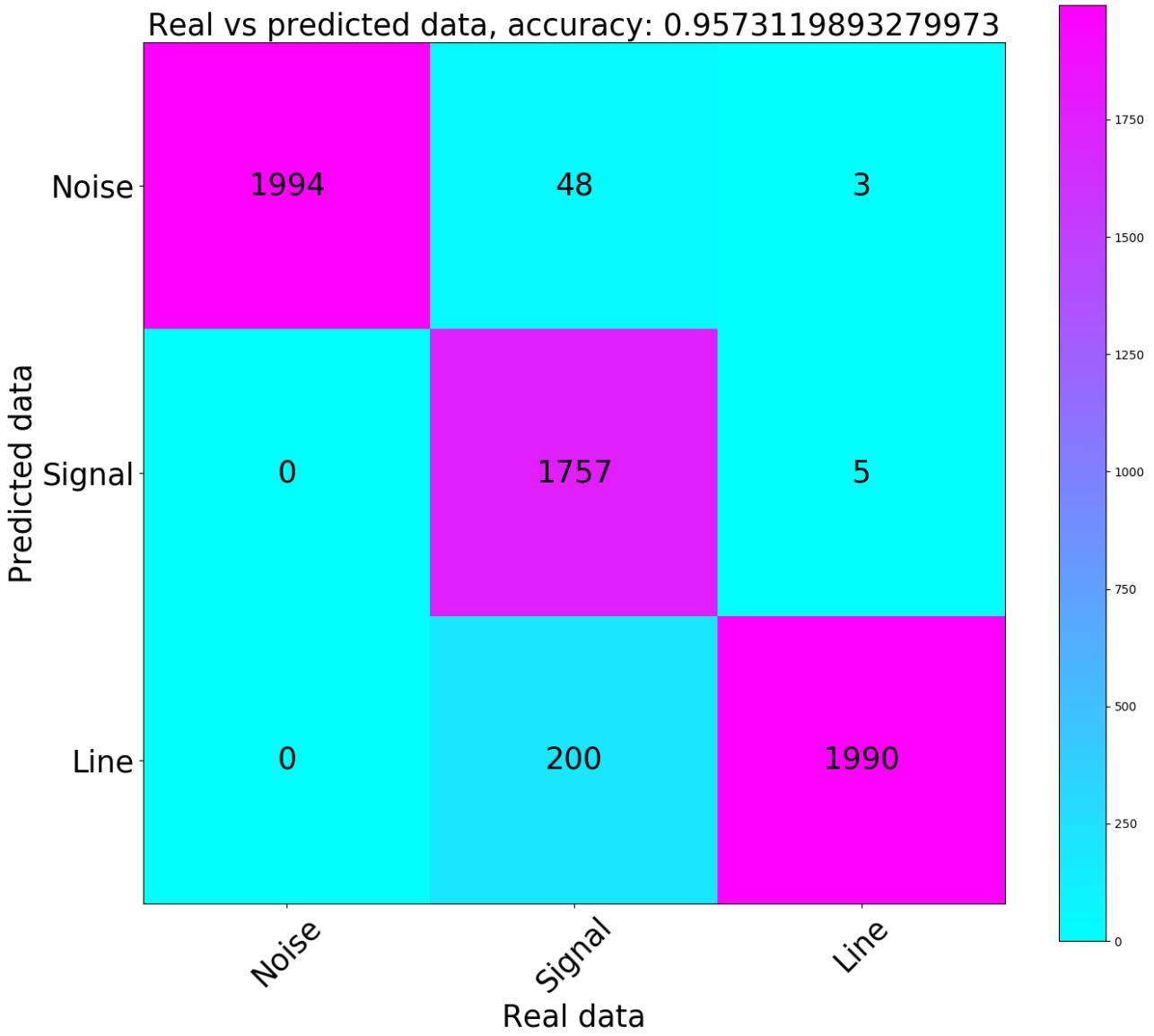


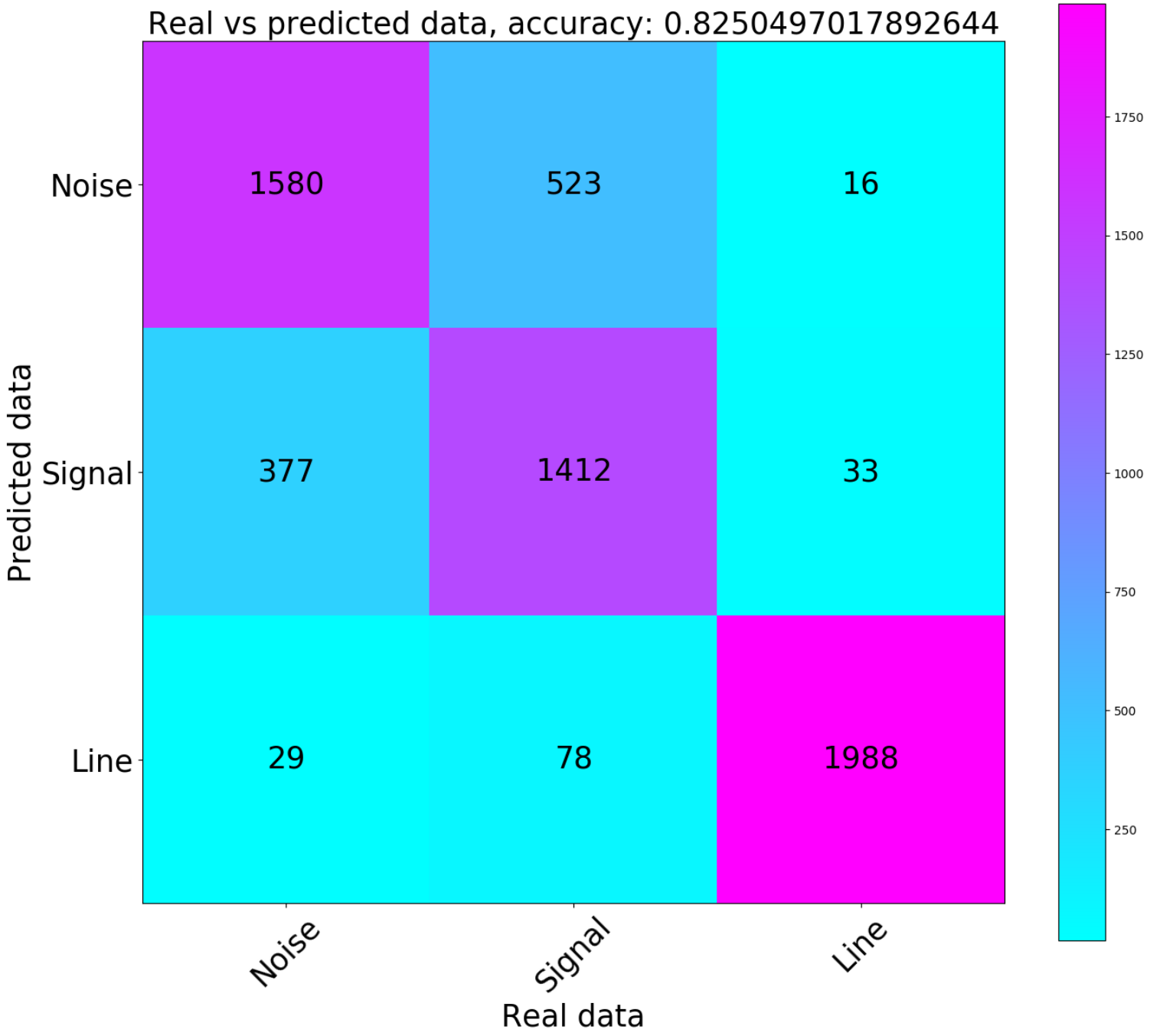
Convolutional neural network - Multi Instance Learning



Common network

source: deeplearning4j.com

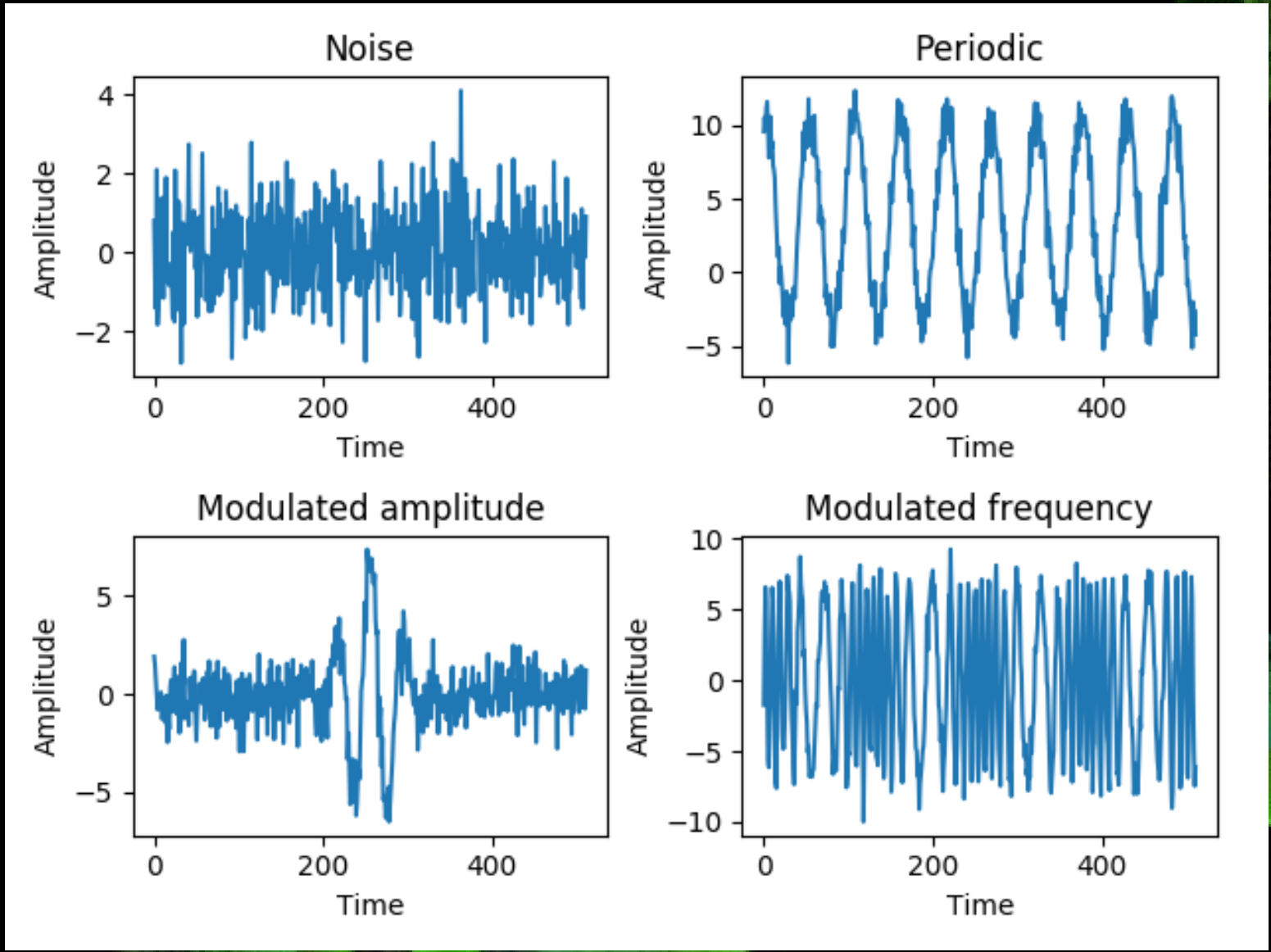

Real vs predicted data, accuracy: 0.9573119893279973
Noise
Signal
Line
Predicted data
1994
48
3
0
1757
5
0
200
1990
Noise
Signal
Line
Real data
1750
1500
1250
1000
750
500
250
0

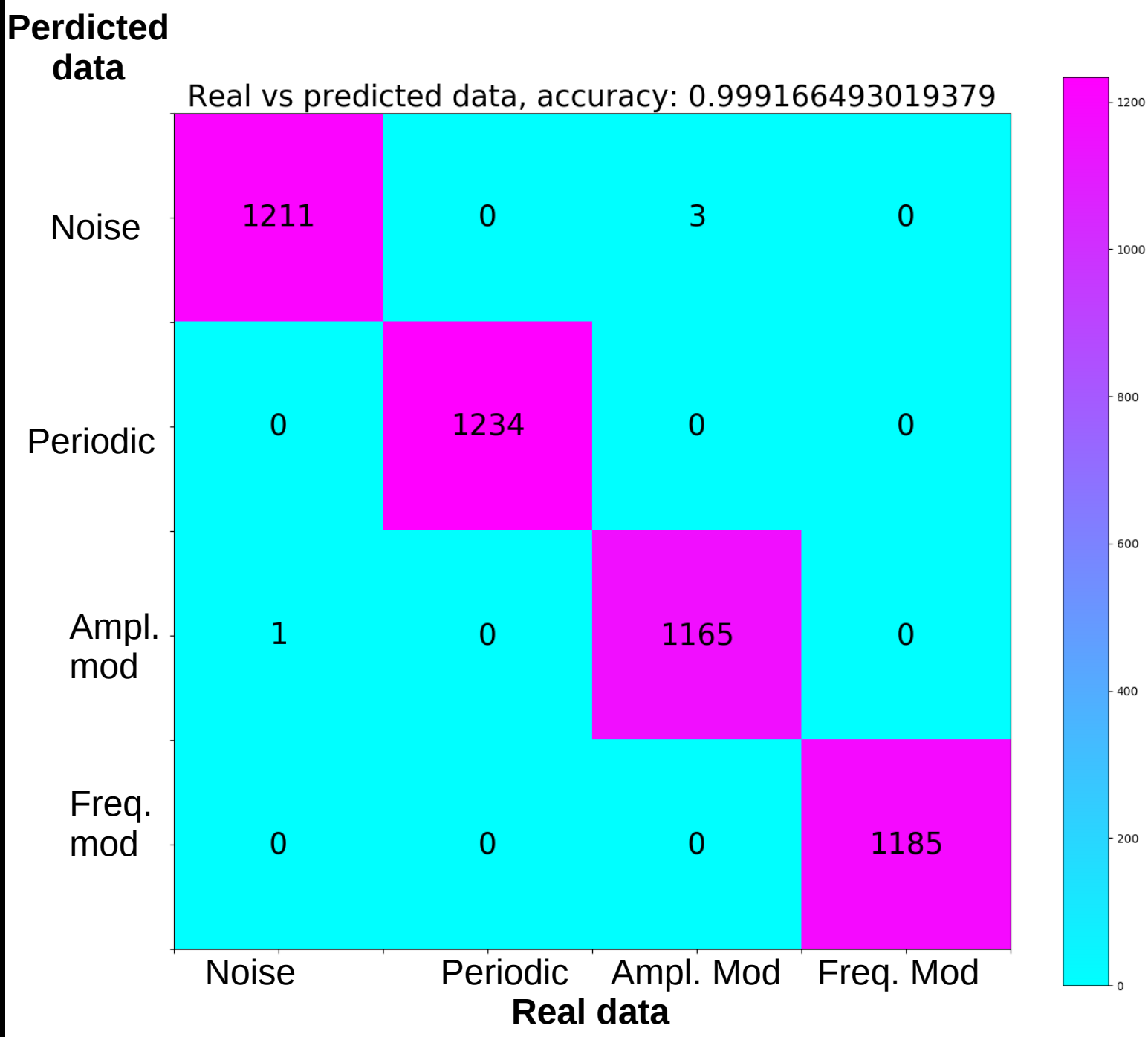

Real vs predicted data, accuracy: 0.8250497017892644
Noise
Signal
Line
1580
523
16
377
1412
33
29
78
1988
Predicted data
Noise
Signal
Line
Real data
1750
1500
1250
1000
750
500
250

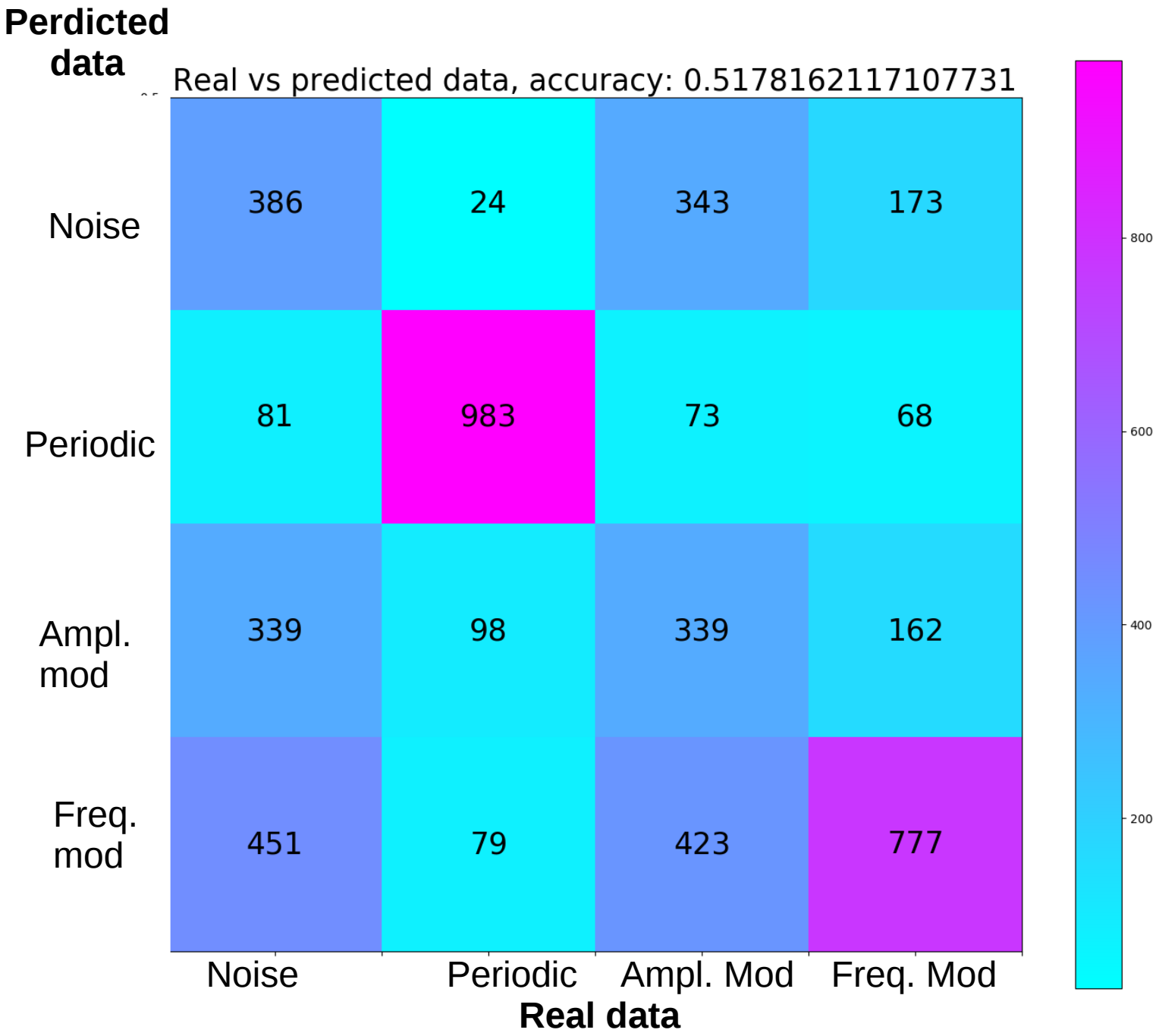
Periodic signal classification

- Long-Short-Term-Memory network – search for patterns in time series data
- Different types of periodicity of the signal

Generated data




Perdicted data
Real vs predicted data, accuracy: 0.999166493019379
Noise
Periodic
Ampl. mod
Freq. mod
1211
0
3
0
0
1234
0
0
1
0
1165
0
0
0
0
1185
Noise
Periodic
Ampl. Mod
Freq. Mod
Real data
1200
1000
800
600
400
200
0


Perdicted data
Real vs predicted data, accuracy: 0.5178162117107731
Noise
Periodic
Ampl. mod
Freq. mod
386
24
343
173
81
983
73
68
339
98
339
162
451
79
423
777
800
600
400
200
Noise
Periodic
Ampl. Mod
Freq. Mod
Real data

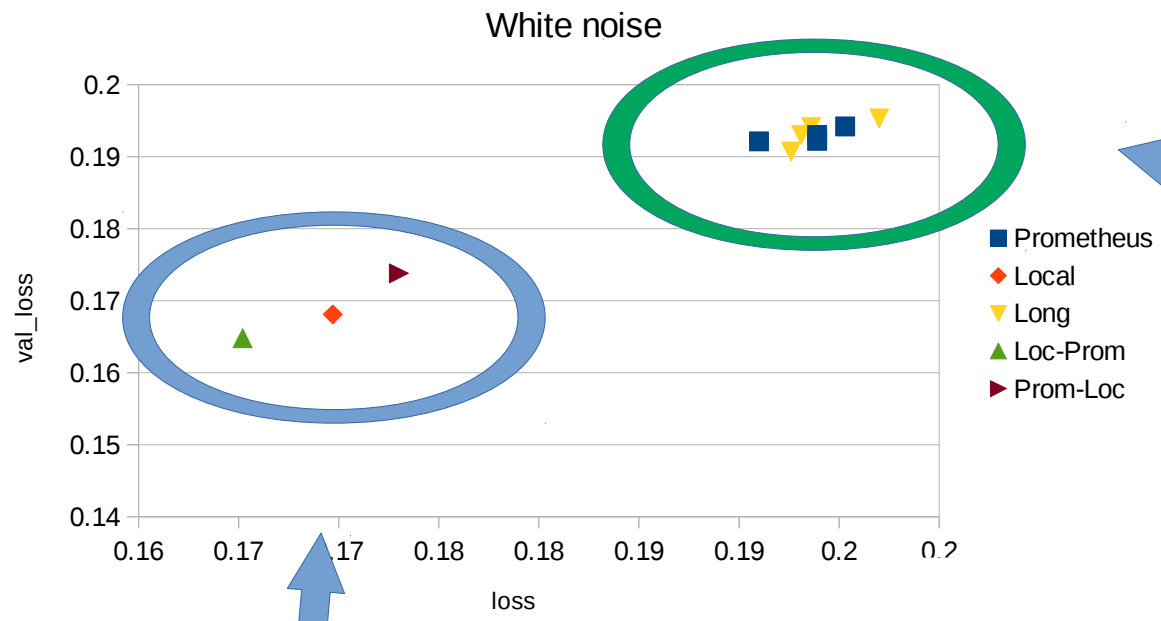
Machine Learning proved to be promising method for signal classification!

Both, in terms of pattern recognition and time-series analysis

Problems

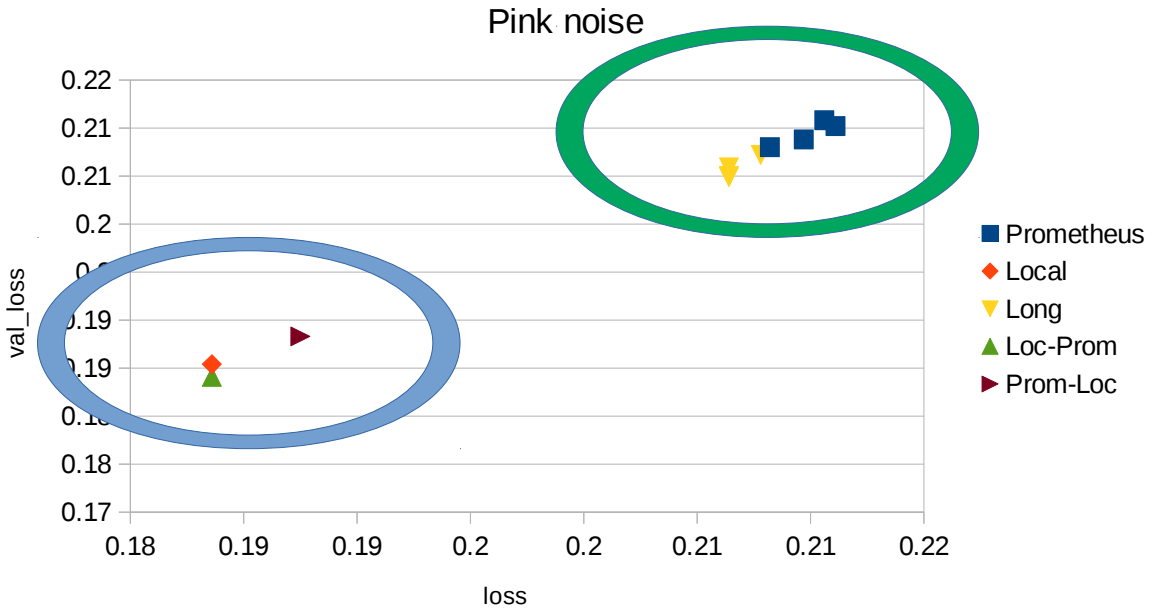
- Performance of computing on CUDA 7 and Tensorflow 1.4
- Parallel GPUs – loss of accuracy; averaging after epochs

Horovod – tests of noise reduction: bigger loss - worse



Parallel GPUs

Single GPU, Mixed GPUs



Future...

- Overtraining issue for high frequency solved
- Merge LSTM and CNN in one MIL
- Test child algorithm
- Test models on multi GPU

Machine Learning proved to be promising method for signal classification!

Both, in terms of pattern recognition and time-series analysis

Thank you for attention!