Applications of machine learning for gravitational wave astrophysics

Chris Messenger - University of Glasgow
Overview

• Brief introduction
• Continuous waves
  • Modelled searches
  • Unmodelled searches
• Burst searches
• Compact binaries
  • Modelled searches
  • Parameter estimation
• Conclusions

PLEASE ASK QUESTIONS THROUGHOUT

https://xkcd.com/1838/
Brief Introduction

Just my opinions
My thoughts (1)

- I’m going to be talking about using machine learning or artificial intelligence or neural networks or … for **astrophysical** gravitational wave data analysis.

- I will be discussing applications towards the existing problems of searching and then characterising astrophysical signals.

- It is heavily biased towards activities at Glasgow (sorry).

- My aim is to highlight the astrophysical problems that need solutions (and to show what we’re doing right now).

- **The black box criticism**
My thoughts (2)

• Why are we doing this?

• Speed
  • Front loading the cost - leads to rapid realtime analysis
  • Simply saves CPU cycles and therefore cash
  • Allows us to keep up with the detections

• It’s just better (maybe)
  • Existing techniques are (near) optimal when you know your model
  • ML has the potential to learn what we have approximated or incorrectly assumed.
  • If faster then computationally limited searches - can be made more sensitive
Continuous waves

Weak but always there
Continuous wave machine learning

- Existing challenges and signal characteristics
  - Vast parameter space
  - Likely very weak signals
  - Narrow band
  - Leading to traditional searches optimised at fixed computational cost - Generally slow

- The latest ML literature
Modelled searches - Dreissigacker et al 2019

- Based on the success of CNNs for compact binary searches
- The task is significantly more difficult here
- Fair comparison with fully coherent searches over a broad parameter space
- The ML approach is reasonably competitive for the simplest of the cases studied
- For $10^6$ sec observations at 1kHz perform significantly worse than matched filtering
- However …

Unmodelled searches - Bayley et al 2019

- A very weakly modelled search for weak psuedo-sinusoidal continuous signals
- Uses the Viterbi algorithm to efficiently find the maximum sum of power/statistic across a time-frequency plane (hence SOAP)
- Requires no templates and runs on “raw” GW data
- Is exceptionally good at finding detector line features (also annoying)
- Extension work applies a CNN to the output for better signal vs line discrimination

Continuous wave challenges

- The parameter space is incredibly large
- The signal is incredibly weak - orders of magnitude lower than the noise amplitude
- The dataset is (quite) large, 1 year X 1 kHz ≈ 10 GB
- In the era of open data the LVC and competitors are keen to analyse the data very quickly
- However, since we have been limited until now by computational expense - with ML this could no longer be a limit, and hence sensitivity can really improve
Burst searches

An ideal ML problem?
Burst analysis machine learning - McGinn et al in prep

- Uses Generative Adversarial network (GAN) to learn how to make standard burst waveforms
- Generation stage has the possibility to make signals spanning all training classes
- Discriminator stage has the potential to be a general transient detection tool

---

12

we wish to test is the time delay so it may be used as a detection flag. First, $D$ was tested by injecting signals with a non-physical time delay of 0.1s and comparing that with the same signal at realistic delays.

Figure 11: The response of $D$ when attempting to classify sine-Gaussian signals with an unphysical time delay between detections, referred to here as a glitch. A sample sine-Gaussian burst simulated using Eq. 3.1 with a physical time delay (top left), the same sample with a 0.1s delay (top right).

$D$’s output to 1000 injects of different signals with physical (purple) and 0.1s delays (orange) (bottom). Fig. D2 shows that there is some indication that the network has learned this parameter as the distributions are not completely overlapping. $D$ classifies the physical signals to be closer to its target of zero, which can be interpreted as being more realistic than the glitch signals. By far the important parameters that the network learns is the shape, staring epoch and frequency. Results for ring-downs and white noise burst can be seen in Appendix D.

One problem with training on burst data set is that there are numerous null values that the network has to learn. It is therefore unsurprising that $D$ classifies these glitches to some degree.

5. Conclusion

This report discussed how machine learning, in particular, generative adversarial networks, may be constructed to search for gravitational wave bursts. Burst signals were represented by three classes: sine-Gaussian, ring-down and white noise bursts. The signals contained an exact copy of itself delayed by a time that was calculated using a randomised sky location relative to LIGO detectors. This aims simulate the output of both detectors and use the time delay as a detection parameter. A GAN was designed through the tuning of hyperparameters that generates noise-free burst signals and outputs a detection statistic that classifies its inputs as bursts. After training, this so-called BurstGAN was then tested on noise realisations of the three classes.
Compact Binaries

These are the ones we’ve detected
Compact binary machine learning

- Existing matched filtering searches are close to optimal and quite fast (~1 min latency)
  - Matched filtering is very close to optimal sensitivity
  - Not all of the parameter space is covered
  - Non-Gaussian noise is well understood

- ML searches can help if they are fast and do as well on detector noise artefacts

- The latest ML literature
CNNs for binary black hole searches - Gabbard et al 2017

- A supervised classification problem.
- Uses a basic and standard CNN network to learn to classify between noise and signal+noise classes.
- Conclusion - the CNN approach can achieve the same sensitivity as a matched filtering analysis.

<table>
<thead>
<tr>
<th>Parameter (Option)</th>
<th>Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
<tr>
<td>Type</td>
<td>C C C C C C H H H</td>
</tr>
<tr>
<td>No. Neurons</td>
<td>8 8 16 16 32 32 64 64 2</td>
</tr>
<tr>
<td>Filter Size</td>
<td>64 32 32 16 16 16 n/a n/a n/a</td>
</tr>
<tr>
<td>MaxPool Size</td>
<td>n/a 8 n/a 6 n/a 4 n/a n/a n/a</td>
</tr>
<tr>
<td>Drop out</td>
<td>0 0 0 0 0 0 0.5 0.5 0</td>
</tr>
<tr>
<td>Act. Func.</td>
<td>Elu Elu Elu Elu Elu Elu Elu Elu SMax</td>
</tr>
</tbody>
</table>

Bayesian parameter estimation
- Gabbard et al 2019

- The current “holy grail” of machine learning for GWs
- We use a variation inference approach to produce samples from the posterior
- It does NOT need to be trained on precomputed posteriors
- It is ~6 orders of magnitude faster than existing sampling techniques
- The paper will be submitted to arXiv this week

Gabbard et al, “Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy”, in prep (2019)
Compact binary challenges

• Binary black holes are easy but binary neutron stars are hard - they are longer in duration and broader in bandwidth

• Spin (with precession) and eccentricity expand the parameter space

• Point estimate parameter estimation is useless - sorry to be harsh.

• Do we even need searches if we can do online PE?
  • Yes, kind of.

• The background estimation problem - a long standing issue that people feel quite strongly about
Conclusions

What I want you to remember
Conclusions

• I have not covered all of the current ML applications to astrophysical GW data analysis - there is a lot going on and not just for ground based detectors

• There are a lot of concept papers but no practical pipelines (LVC perspective)

• I have also outlined only some of the problems we have

• For me, it seems completely obvious that all data analysis will be ML in 5-10 years

• Back to the **black box**
Many thanks