Application of Machine Learning techniques to the discovery of new physics

Óliver Partida Gutiérrez

UAB

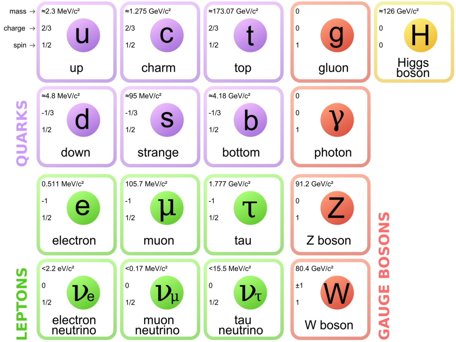
November 8, 2020

Outline

1. Problem description
2. Machine Learning techniques
3. Results
4. Conclusions

Flavor-changing neutral currents FCNC

- $B \to K^* \mu\mu$
- $\Delta S = \Delta Q = \pm 1$.
- $\Delta S = 2$ and FCNC only at second order in SM.
- Very suppressed processes by GIM mechanism.
- Candidates to search for new physics.



GIM mechanism

- Put forth by Sheldon Lee Glashow, John Iliopoulos and Luciano Maiani.
- Required the introduction of a new quark.
- mass-squared difference of the different virtual quarks exchanged in the box diagram.

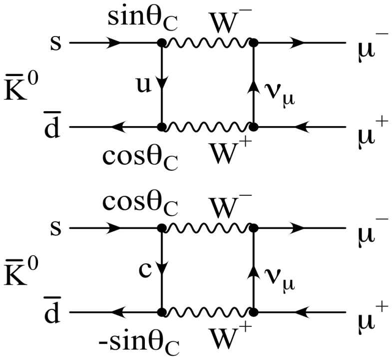


Figure:

Neutral Kaon decay [17]

Experimental data

- In 2013, the LHCb collaboration announced the measurement of angular observables describing the decay $B \to K\mu\mu$.
- Two observables, P_2 and P_5' were in significant disagreement with the SM expectations.
- $RK = Br(B \to K\mu\mu)/Br(B \to Kee)$ measurement $\neq 1$. Could indicate Lepton Flavor Universality Violation (LFUV).

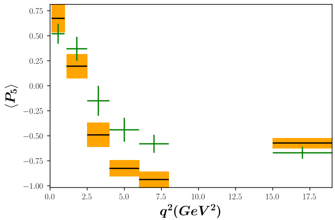


Figure: Observable P_5.

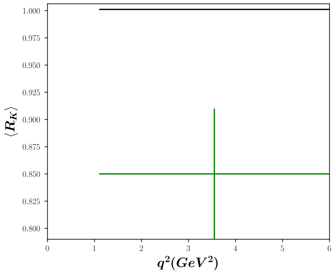


Figure: $R_K = Br(B \to K\mu\mu)/Br(B \to Kee)$.

Searching for New Physics

- Effective field theory framework (EFT)

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_i \mathcal{C}_i \mathcal{O}_i.$$

- Heavy degrees of freedom (the top quark, the W and Z bosons, the Higgs and any heavy new particle) are integrated out in short-distance Wilson coefficients $\mathcal{C}_i$, leaving only a set of operators $\mathcal{O}_i$ describing the physics at long distances.
- In the SM, the Hamiltonian contains 10 main operators with specific chiralities due to the V - A structure of the weak interactions.
- We reduce such set to the two dominant operators:

$$\mathcal{O}_9 = \frac{e}{16\pi^2} m_b (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \ell),$$

$$\mathcal{O}_{10} = \frac{e}{16\pi^2} m_b (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \gamma_5 \ell).$$

Searching for New Physics

- Current analysis of anomalies in flavor physics are based on a linear regression of a χ^2 function.
- The observables measured by the experimental collaborations can be written in terms of the C_i Wilson coefficients.
- It is common to split the Wilson coefficients in two pieces: the SM contributions and the NP contributions:

$$C_9 = C_9^{SM} + C_9^{NP}$$

$$C_{10} = C_{10}^{SM} + C_{10}^{NP}$$

Discriminative vs Generative

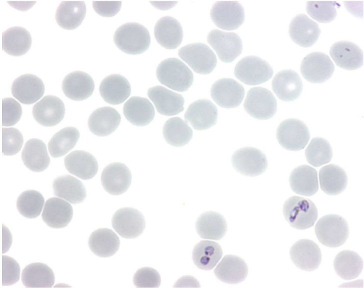
- Two distinct approaches to solving decision problems.
- Generative:
  - Determine the class-conditional densities $p(x|C_k)$ for each class individually and the prior class probabilities.
  - Use Bayes theorem to find the posterior:

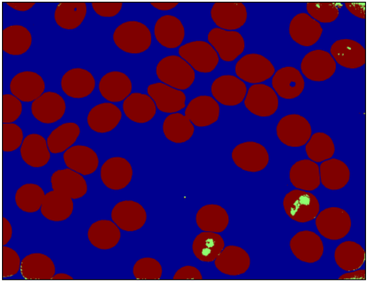
$$p(C_k|x) = \frac{p(x|C_k)p(C_k)}{p(x)}$$

  - We have a model of our data. We can sample from it.
  - E.g.: Naive Bayes, Mixture Models, Generative Adversarial Networks.
- Discriminative:
  - We have labeled data.
  - Directly determine the posterior class probabilities $p(C_k|x)$.
  - E.g.: Neural Networks.

Generative models

- $p(x|\mu, \Sigma, \pi) = \sum_{i=1}^{k} \pi_i p(x|\mu_i, \Sigma_i)$.
- Model the probability distribution of the data.
- Fit model parameters using Maximum Likelihood.
- Use Bayes theorem to classify new data point.





Generative Adversarial Networks (GANs)

- First described in the 2014 paper by Ian Goodfellow [16].
- **What are they?** $\rightarrow$ Model architecture for training a generative model.
- Implicitly learn rich distributions hard to model with an explicit likelihood.

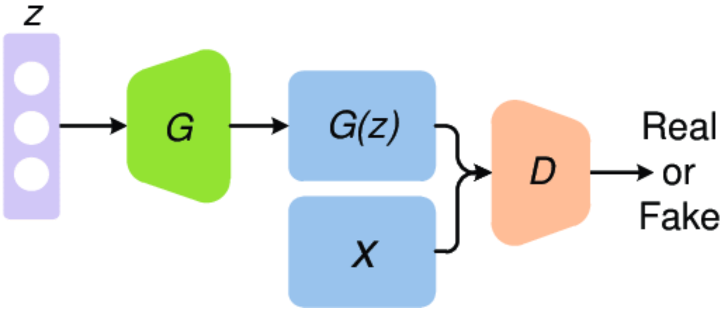


Figure: **GAN architecture**, from Recent Progress on Generative Adversarial Networks (GANs): A Survey [15]

GAN training problems

- **Vanishing gradients** $\rightarrow$ optimal discriminator doesn't provide enough information for the generator to make progress.
- **Mode collapse** $\rightarrow$ if a generator produces an especially plausible output, the generator may learn to produce only that output.
- **Failure to converge** $\rightarrow$ If the generator succeeds perfectly, then the discriminator has a 50% accuracy and the discriminator feedback gets less meaningful over time.

Results. Dataset

- 8 observables.
- 37 bin values.
- 10000 samples.
- 10×10 $C_9 \in [-2, 0]$, $C_{10} \in [-1, 1]$.
- 100 $a_0, a_1, a_2 \in [-0.01, 0.01]$.

$C_9 = -1.33$, $C_{10} = -0.44$ $\chi^2 = 23.33$	a_0	a_1	a_2
P_1	-0.093	0.001	0
P_2	-0.093	0.001	0
P_4	-0.093	0.001	0
P_5	-0.093	0.001	0
BrK^{0*}	-0.093	0.001	0
BrK^{0}	-0.093	0.001	0
R_K	-0.093	0.001	0
R_{K^*}	-0.093	0.001	0

Figure: Coefficients of sample in training set with lowest χ^2 value.

Results. Neural networks

$C_9 = -0.79,\ C_{10} = -0.07,\ \chi^2 = 29.5$	a_0	a_1	a_2
P_1	0	0	0
P_2	0	0	0
P_4	0.03	0	0
P_5	0.02	0	0
BrK^{0*}	-0.01	0	0
BrK^0	-0.01	0	0
R_K	-0.01	0	0
R_{K^*}	-0.01	0	0

Figure: Coefficients predicted by the NN after training

Results. GAN

$C_9 = -0.74,\ C_{10} = -0.62\ ,\ \chi^2 = 373.48$	a_0	a_1	a_2
P_1	0.012	0.016	0.0008
P_2	-0.017	0.030	-0.002
P_4	-0.010	0.010	-0.001
P_5	0.001	0.008	0.0003
BrK^{0*}	-0.026	0.023	0.001
BrK^{0}	-0.008	0.014	0.005
R_K	-0.019	0.012	-0.004
R_{K^*}	-0.013	0.011	-0.006

Figure: Coefficients with lowest χ^2 sampled from GAN.

Conclusions

- The parametrization of the observables

Thank you

- M. Algueró, B. Capdevila, A. Crivellin, S. Descotes-Genon, P. Masjuan, J. Matias, M. Novoa and J. Virto, Eur. Phys. J. C **79** (2019) no.8, 714 [arXiv:1903.09578 [hep-ph]].
- M. Algueró, B. Capdevila, S. Descotes-Genon, P. Masjuan and J. Matias, Phys. Rev. D **99** (2019) no.7, 075017 [arXiv:1809.08447 [hep-ph]].
- A. Vicente, [arXiv:2001.04788 [hep-ph]].
- R. Aaij *et al.* [LHCb Collaboration], arXiv:1903.09252 [hep-ex].
- S. Choudhury for the Belle collaboration, "Measurement of Lepton Flavour Universality in B decays at Belle", Talk at 'EPS-HEP Conference 2019'.
- R. Aaij *et al.* [LHCb Collaboration], JHEP **1406** (2014) 133 [arXiv:1403.8044 [hep-ex]].
- A. Abdesselam *et al.* [Belle Collaboration], arXiv:1904.02440 [hep-ex].
- M. Aaboud *et al.* [ATLAS Collaboration], [arXiv:1812.03017 [hep-ex]].

- S. Aoki *et al.* [Flavour Lattice Averaging Group], arXiv:1902.08191 [hep-lat].
- S. Wehle *et al.* [Belle Collaboration], Phys. Rev. Lett. **118** (2017) no.11, 111801 [arXiv:1612.05014 [hep-ex]].
- A. Abdesselam *et al.* [Belle Collaboration], arXiv:1604.04042 [hep-ex].
- L. Calibbi, A. Crivellin and T. Li, Phys. Rev. D **98** (2018) no.11, 115002 [arXiv:1709.00692 [hep-ph]].
- A. Crivellin, D. Muller and C. Wiegand, arXiv:1903.10440 [hep-ph].
- C. Bobeth, A. J. Buras, A. Celis and M. Jung, JHEP **1704** (2017) 079 [arXiv:1609.04783 [hep-ph]].
- Z. Pan, W. Yu, X. Yi, A. Khan, F. Yuan and Y. Zheng, "Recent Progress on Generative Adversarial Networks (GANs): A Survey," in IEEE Access, vol. 7, pp. 36322-36333, 2019, doi: 10.1109/ACCESS.2019.2905015.

- Ian J. Goodfellow and Jean Pouget-Abadie and Mehdi Mirza and Bing Xu and David Warde-Farley and Sherjil Ozair and Aaron Courville and Yoshua Bengio [arXiv:1406.2661 [stat.ML]].
- GIM mechanism — Wikipedia, The Free Encyclopedia, `https://en.wikipedia.org/w/index.php?title=GIM_mechanism&oldid=958923133`