Machine Learning with $D^0 \to K_s^0\pi^+\pi^-$ at LHCb

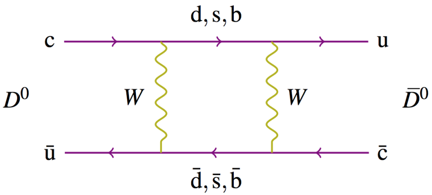
Martha Hilton

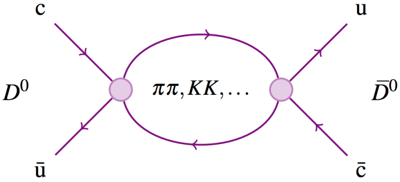
martha.hilton@cern.ch

University of Manchester

January 8, 2019

$D^0 \to K_s^0 \pi^+ \pi^-$





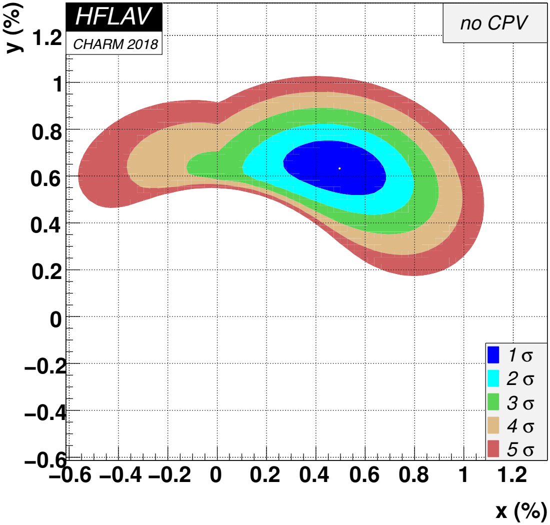
Mixing parameters:

$$x \equiv \frac{(m_1 - m_2)}{\Gamma} \qquad y \equiv \frac{(\Gamma_1 - \Gamma_2)}{2\Gamma}$$

Mass Eigenstates:

$$\left|D_{1,2}\right\rangle = p\left|D^0\right\rangle \pm q\left|\bar{D}^0\right\rangle$$

World Average



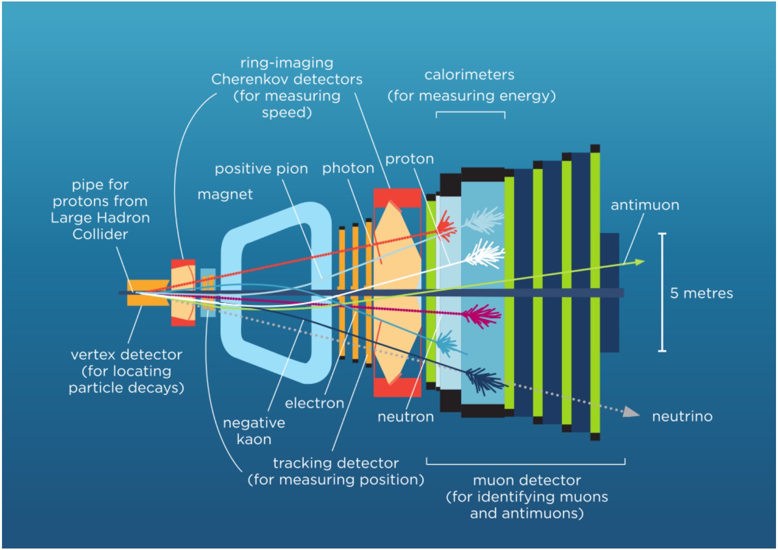
Neutral Meson Mixing

$$\left|\mathcal{A}_f(t)\right|^2 =$$

$$\frac{1}{2}e^{-\Gamma t}\left[\left(|\mathcal{A}_f|^2 - \left|\frac{q}{p}\bar{\mathcal{A}}_f\right|^2\right)\cos(x\Gamma t) - 2\mathcal{I}m\left(\mathcal{A}_f\bar{\mathcal{A}}_f^*\left[\frac{q}{p}\right]^*\right)\sin(x\Gamma t)\right.$$

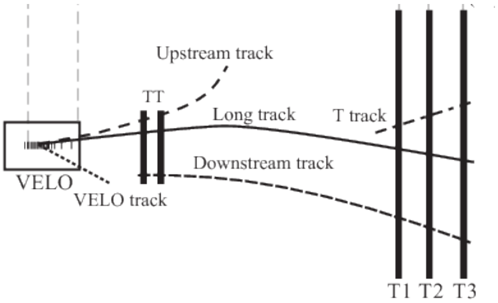
$$\left. + \left(|\mathcal{A}_f|^2 + \left|\frac{q}{p}\bar{\mathcal{A}}_f\right|^2\right)\cosh(y\Gamma t) - 2\mathcal{R}e\left(\mathcal{A}_f\bar{\mathcal{A}}_f^*\left[\frac{q}{p}\right]^*\right)\sinh(y\Gamma t)\right]$$

The LHCb Detector



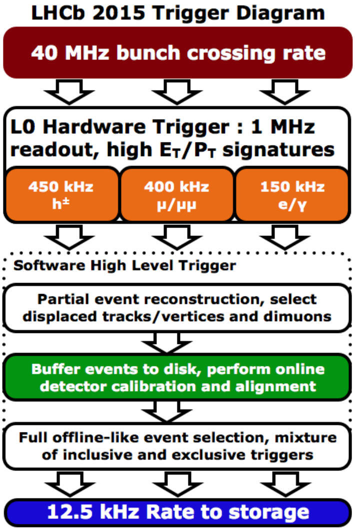
Data samples

- **Single-tagged:** $B^- \to D^0(\to K_s^0\pi^+\pi^-)\mu^-\bar{\nu_\mu}$
- **Double-tagged:** $\bar{B}^0 \to D^{*+}(\to D^0(\to K_s^0\pi^+\pi^-)\pi^+)\mu^-\bar{\nu_\mu}$



- **VELO:** Vertex Locator
- **TT and T1-T3:** Tracking stations located before and after the dipole magnet

The LHCb Data Flow



Pre-Selection

Pre-selection Requirements

- B meson flight distance χ^2: $\chi^2_{FD} > 100$
- $\boldsymbol{\chi^2_{FD}}$ **:** Difference between the χ^2 of the primary vertex fit with and without the particle
- Probability that the muon is a muon: $\textsc{ProbNNmu} \geq 0.6$
- Pion does not look like a muon

Decay Tree Fitter

- Kinematic refit with known D^0 and K^0_s masses
- Fit must have converged
- Fit quality must be good: $\chi^2 < 25$

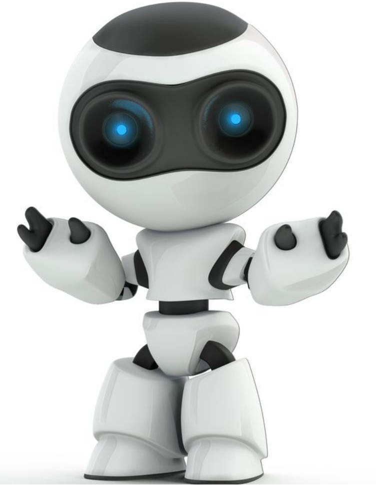
What is Machine Learning?

Types of Machine Learning:

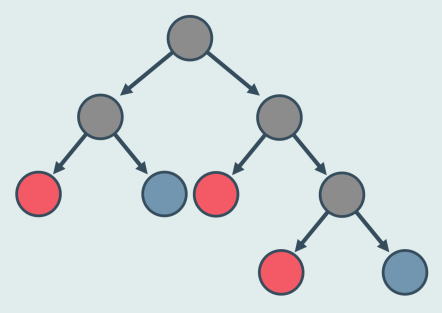
- **Supervised:** Data points have known outcome
- **Unsupervised:** Data points have unknown outcome

Types of Supervised Learning:

- **Regression:** Predicted quantity is continuous
- **Classification:** Predicted quantity is a category



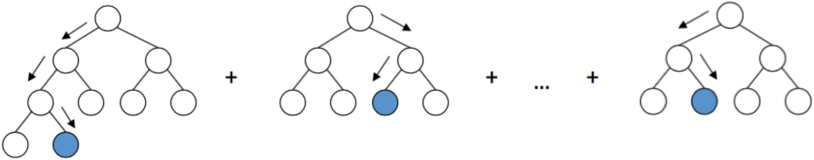
Boosted Decision Trees



- **Features:** properties of the data used for prediction
- **Target:** predicted category or label of the data
- **Label:** the target value for a single data point

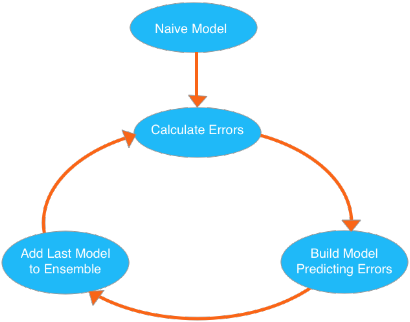
Gradient Boosted Decision Trees

- Gradient Boosting is an ensembling technique
- Prediction is done by an ensemble of estimators
- Gradient boosting provides improved predictions by combining an ensemble of estimators
- **Weak Learner:** A decision tree which does not perform very well



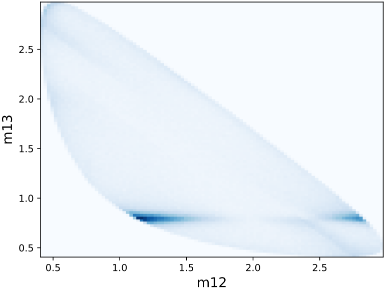
XGBoost

- **XGBoost:** e**X**treme **G**radient **Boost**ing
- Can be used in a range of computing environments: parallelisation and distributed computing
- XGBoost is much faster than other gradient boosting implementations



Boosting to Uniformity

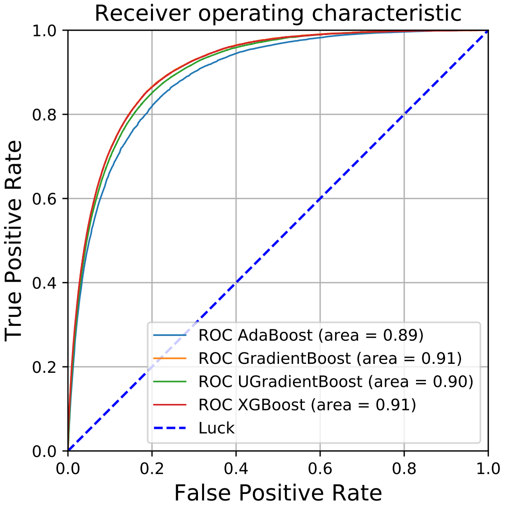
- It is important to have flat efficiency in decay time and Dalitz variables to avoid biases
- uBoost uses a combination of FlatnessLoss and AdaLoss
- **FlatnessLoss:** penalises non-uniformity
- **AdaLoss:** penalises poor predictions like regular boosting
- **Loss Function:** $\mathcal{L} = \mathcal{L}_{AdaLoss} + \alpha\mathcal{L}_{FL}$



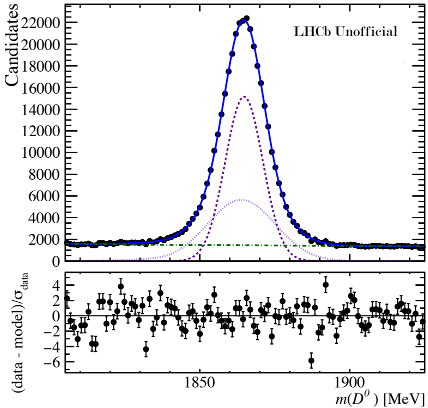
BDT Input Variables

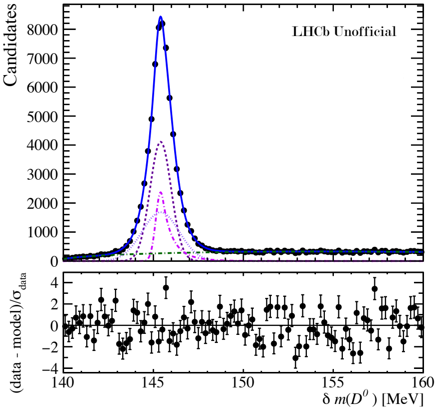
- $m_{corr}(B)$: corrected mass of the B-meson candidate with respect to the best primary vertex
- **Corrected Mass:** Mass corrected for the missing transverse energy of the neutrino
- θ_{DIRA} of the B-meson candidate
- **DIRA (Direction Angle)**: The angle between a line drawn from the primary vertex to the decay vertex of the particle and the sum of the 4-momentum of its decay products
- $\chi^2/ndof$ of the B-meson decay vertex
- χ^2 of the B-meson primary vertex
- p_T of the muon candidate
- p_T of the D^0 candidate
- $\chi^2/ndof$ of the D^0 meson decay vertex

Multivariate Analysis



Mass Plots





Delta mass: $\delta m = m(D^*) - m(D^0)$

Summary

- Measure mixing parameters of D^0 meson using $D^0 \to K_s^0\pi^+\pi^-$ decay at LHCb detector
- Mixing is suppressed in the charm sector and experimentally challenging
- $D^0 \to K_s^0\pi^+\pi^-$ offers direct access to the mixing parameters x and y
- Machine learning is crucial for the amplitude fit used in this analysis