# Notes NOC

ALL

* Allele sharing
* Low-template DNA (low quality) -> drop-out
* Elevated stutter/drop-in

## nC-tool [1]

Derive NOC from

* TAC
* Reduce nr. of alleles by 5% from expected drop-out/alle sharing

## RFC19 model [2]

Derive NOC from

* Random forest ML model with 19 features

Percentage correct predictions for 2-5 NOC

* MAC = 69.2%
* nC-tool = 76.7%
* RFC19 model = 85%

Output = NOC + prob

## Decision Tree [3]

Derive NOC from

* Decision tree with

Tested various ML approached (RF / MLP / LDA), showing similar performance to the RF19 model. They obtained very high performance (96%) with a RFC 35 model.

Difference with [2] is “Benschop et al. used 1174 unique donors to construct 590 profiles [20], whereas the PROVEDIt dataset only had 26 unique donors within the 766 profiles used”

This means that the classifiers probably overfit to certain donors.

“In conclusion, the decision tree method for NoC assignment has been shown to be over 77% accurate, with increasing performance with improved stutter and artefact filters”

They used a decision tree to classify peaks as stutter or allele.

## Background STR mixture interpretation [4]

Information about how statistical analysis is done to determine the LR with the Hd and Hp. Showing that the LR is still the de-facto standard method.

“The peak height information is of benefit for analyzing mixed profiles.”

“The effect of incorrect estimation of the number of donors (caused by allele sharing) to the LR value was examined by Benschop (…) and was illustrated to exert a great effect on the LR” [5]

## Background about NFI-used software for LR calculation DNAStatistX

Shows the importance of correct NOC estimations: under-assigned number of contributors can cause the model the fail calculating the LR because the observed peaks cannot be well explained.

Also includes the NOC model + the generic RF11 model (with a lower accuracy of ~

Also includes the LoCIM method for inferring the major contributor.

## NOC

Only using the MAC, can severely impact performance, [6], more than 70% of 4-person mixtures were characterized as 2-, or 3-person mixtures.

Allowing stutter peaks to be counted as alleles [7].

MAC in comparison to Maximum Likelihood [8]

How relatives influence the LR [9]

General rules for interpreting mixtures [10]: “1) the reasoning used to define the sample as being a mixture, (2) attempting to define the number of contributors to the mixture, (3) attempting to determine the ratio of the contributors to define major or minor genetic profiles or to deconvolute mixtures, (4) taking stochastic thresholds into consideration to determine which loci may have incomplete reportable DNA profiles, and (5) taking peak height ratios of the components into consideration to determine which peaks must be designated as indistinguishable from stutter.”

“LR varied considerably when the hypotheses used an incorrect number of contributors” [5]

More contributors, more likely to be estimated to have fewer NOC [11].

# Notes metrics

* Accurate w.r.t. true effects of variables
* Speed?

## Counterfactuals [12]

“MOC returns a Pareto set of counterfactuals that represents different trade-offs between our proposed objectives, and which are constructed to be diverse in feature space.”

* Low number of feature changes (sparse explanations)
* Close to nearest observed data points (plausible explanations)

## ICE / PD plots /

ICE paper about PD plots: “Note that the approximation here is twofold: we estimate the true model with fˆ, the output of a statistical learning algorithm, and we estimate the integral over xC by averaging over the N xC values observed in the training set.”

“Visually, ICE plots disaggregate the output of classical PDPs. Rather than plot the target covariates’ average partial effect on the predicted response, we instead plot the N estimated conditional expectation curves: each reflects the predicted response as a function of covariate xS, conditional on an observed xC.”

Dependence between features must be visualized in explanations

The quality of explanations is sometimes evaluated by performing a quantitative evaluation of a user study. Users are asked to perform a certain task and the explanations help support this task. How well and how fast the humans can accomplish the task is measured as accuracy and efficiency respectively [13, 14]. Subjectively, users were asked for their preference of explanation type in a 1 versus 1 fashion and asked to provide reasons.

Explanations should have few features, as humans pick usually just a few reasons. They should be specific to the problem at hand, and every instance should be explained in the same deterministic way [15]. Deterministic, or consistent feature attributions [16].

Exploration using several visual aids [17].

The literature about explainable machine learning on tabular data often refers to domains where such tools are used in decision support settings. Domain experts are provided with an ML model which has automated their task such as determining whether or not someone is granted a loan, or whether a patient is at risk for developing cancer. The explanations are then required to help the experts determine if the prediction is trustworthy. Sometimes these explanations do not seem to correspond well with intuition. This can be caused by various underlying issues, but this is not often made clear to the user. Examples of such issues are

* The features listed should not contribute to the prediction in the way that they are shown
  + Maybe the feature is highly correlated with another feature, which causes the model to assign all contribution to one, therefore skewing perception.
  + Maybe the feature value is underrepresented in the training set, which causes the model to make decisions on little data, with as a result, a poor generalization
* The model is uncertain

Show prototypes and criticisms

Research into XAI has shown the need for comparison and evaluation of methods [18-24], and the recent interest in the implementation of counterfactual explanations [18-22]. Although there are a few key components hightlighted by these surveys, they also mention that the evaluation must be done specifically to certain applications [22]. One could specify a specific goal to be achieved by the explanations which should be tested [23]. Also the relevance of explanations to a certain audience [24].

1. Benschop, C., A. Backx, and T. Sijen, *Automated estimation of the number of contributors in autosomal STR profiles.* Forensic Science International: Genetics Supplement Series, 2019. **7**.

2. Benschop, C.C.G., et al., *Automated estimation of the number of contributors in autosomal short tandem repeat profiles using a machine learning approach.* Forensic Science International: Genetics, 2019. **43**: p. 102150.

3. Kruijver, M., et al., *Estimating the number of contributors to a DNA profile using decision trees.* Forensic Science International: Genetics.

4. Tao, R., et al., *Separation/extraction, detection, and interpretation of DNA mixtures in forensic science (review).* International Journal of Legal Medicine, 2018. **132**.

5. Benschop, C.C.G., et al., *The effect of varying the number of contributors on likelihood ratios for complex DNA mixtures.* Forensic Science International: Genetics, 2015. **19**: p. 92-99.

6. Paoletti, D.R., et al., *Empirical analysis of the STR profiles resulting from conceptual mixtures.* Journal of Forensic Sciences, 2005. **50**(6): p. 1361-1366.

7. Gill, P., et al., *Interpretation of complex DNA profiles using empirical models and a method to measure their robustness.* Forensic Science International: Genetics, 2008. **2**(2): p. 91-103.

8. Haned, H., et al., *Estimating the Number of Contributors to Forensic DNA Mixtures: Does Maximum Likelihood Perform Better Than Maximum Allele Count?* Journal of Forensic Sciences, 2011. **56**(1): p. 23-28.

9. Dørum, G. and T. Egeland, *Likelihood ratios for complex mixtures with relatives.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e61-e62.

10. Heidebrecht, B.J., *Mixture Interpretation (Interpretation of Mixed DNA Profiles with STRs Only)*, in *Encyclopedia of Forensic Sciences: Second Edition*. 2013. p. 243-251.

11. Coble, M.D., et al., *Uncertainty in the number of contributors in the proposed new CODIS set.* Forensic Science International: Genetics, 2015. **19**: p. 207-211.

12. Dandl, S., et al. *Multi-Objective Counterfactual Explanations*. in *Parallel Problem Solving from Nature – PPSN XVI*. 2020. Cham: Springer International Publishing.

13. Kim, B., O. Koyejo, and R. Khanna. *Examples are not enough, learn to criticize! Criticism for Interpretability*. in *NIPS*. 2016.

14. Ribeiro, M.T., S. Singh, and C. Guestrin. *Anchors: High-Precision Model-Agnostic Explanations*. in *AAAI*. 2018.

15. Lakkaraju, H., et al., *Interpretable & Explorable Approximations of Black Box Models.* 2017.

16. Lundberg, S. and S.-I. Lee, *A Unified Approach to Interpreting Model Predictions*. 2017.

17. Lundberg, S.M., et al., *Explainable machine-learning predictions for the prevention of hypoxaemia during surgery.* Nat Biomed Eng, 2018. **2**(10): p. 749-760.

18. Du, M., N. Liu, and X. Hu, *Techniques for interpretable machine learning.* Communications of the ACM, 2020. **63**(1): p. 68-77.

19. Adadi, A. and M. Berrada, *Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI).* IEEE Access, 2018. **6**: p. 52138-52160.

20. Barredo Arrieta, A., et al., *Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI.* Information Fusion, 2020. **58**: p. 82-115.

21. Carvalho, D.V., E.M. Pereira, and J.S. Cardoso, *Machine learning interpretability: A survey on methods and metrics.* Electronics (Switzerland), 2019. **8**(8).

22. Gilpin, L.H., et al. *Explaining Explanations: An Overview of Interpretability of Machine Learning*. in *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*. 2018.

23. Lipton, Z.C., *The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery.* Queue, 2018. **16**(3).

24. Murdoch, W.J., et al., *Definitions, methods, and applications in interpretable machine learning.* Proceedings of the National Academy of Sciences of the United States of America, 2019. **116**(44): p. 22071-22080.