17th International Congress on Artificial Intelligence - August 21-22, 2024 – Türkiye by IKSAD Institute

**AI-based derivation of the importance of attributes in case of evaluation models**

**Gyula Bán** (<https://orcid.org/0009-0007-1769-3930>),

János Rikk (<https://orcid.org/0000-0002-3846-6661>)

László Pitlik (<https://orcid.org/0000-0001-5819-0319>)

e-Mails: gyulaa.ban@gmail.com, rikk.janos@kodolanyi.hu, pitlik@my-x.hu

Kodolányi University and MY-X research team Hungary

*Keywords: anti-discrimination, optimization, automation, similarity analysis, hermeneutics*

**Abstract**

History: The relevance/importance of attributes is a well-known phenomenon on the field of the statistics. However, these interpretations seem to be naïve (not ooptimized/validated) despite of statistical theories. At first, correlation-based human intuitions (c.f. the attributes with higher correlation are more important) needs at least two attributes (and none of them may be monotonous - c.f. production functions). Linearity and or mixed ceteris paribus constellations (see N [kg/ha] vs. yield [kg/ha] in case of maize with different soil-conditions) can be seen as critical aspects of the correlation-based approaches. Let alone: there are unlimited types of patterns, which can not be described with the formulas of the well-known correlations. One of the previous publications about the real impact/value of the attributes within models declared: stones = sand-elements, or even: it is possible to build diamond from CO2! It means: the importance of the attributes is a relative phenomenon. The importance of a particular attribute is depending on the set of the given attributes, because the reality seems to be very complex (each potential attribute does have a relationship with each other ones) and a real set of attributes is always a partial description of the unlimited complex reality. Therefore, it is necessary to search for more accurate approximations in this field (c.f. <https://miau.my-x.hu/miau/274/real_values_of_attributes.docx>).

Background and benchmarks: Classic production functions have always two/more realistic attributes (c.f. Y and Xi). These attributes make possible to calculate the correlation values. On the other hand, the anti-discriminative optimization (c.f. similarity analysis-based evaluation: e.g. COCO Y0 - <https://miau.my-x.hu/miau/196/My-X%20Team_A5%20fuzet_EN_jav.pdf>) does not have two real attributes, because the hypothetical Y0-variable has only a constant value – and so, correlation can not be calculated based on these raw data.

Highlighted details: If an OAM (object-attribute-matrix) is given, where the objects are e.g. human beings and the attribute are their descriptive phenomena, then it can become necessary to derive an aggregated index value (e.g. index of capability to do something by the human individuals – c.f. military actions). Based on these OAMs, anti-discriminative models can be derived, where the output is a kind of estimated index-value (see above). The question is: how may we interpret, if always one single attribute is excluded from the entire OAM and the aggregated index values will be calculated on these partial OAMs? Which kind of changes can be expected concerning the output-values (= aggregated index values)? Based on these potential changes, what kind of attributes seem to be more relevant for the evaluation (for the index-building-processes)? Finally, what is relevance/importance as such?

The results more complex than expected: The classic output version is a ranking position for each object - based on all input-attributes. Therefore, the classic output versions are in case of the reduced OAMs, similar/identical ranking values from 1 to n, where n is the number of the objects. In the basic case, not one single ranking value should be existing in two or more copies. But at least the anti-discriminative evaluation has to prove, whether all objects may have the same evaluation index-value (therefore: the same ranking value). If this scenario is given, then it is to analyse: are all potential attributes involved into the aggregated evaluation or not? In the very complex combinatorial space, attributes can be described based on a lot of specific attributes (like standard deviation among the output-values in case of a given exclusion). If the raw attributes for object evaluation lead to a new OAM, where the previous attributes (the lacks of them as scenarios) are the new objects, then the anti-discriminative modelling deliver (validated) estimation for the importance of the raw attributes (c.f. <https://miau.my-x.hu/miau/314/importance.xlsx>). More complex interpretation will be published in the full text.

Future aspects: After closing the manual-driven test-cases, the entire evaluation process can be automated e.g. in frame of development task for a bachelor's degree following the KNUTH’s principle: knowledge/science is, what can be transformed/transferred/trans-scripted into source codes – each other human activity is an artistic performance!