Iwein Vranckx *October 14, 1978*

Last update on January 1, 2021

ivranckx@gmail.com • 016.43.83.29 • 0475.75.34.92 • Github • Researchgate • Linkedin
Oudenbos 81, 3202 Rillaar, Belgium • 0000-0002-2439-6906

> *"Outliers are like thinking about an enigma. There it is before you - smiling, frowning, inviting, grand, mean, insipid, or savage, and always mute with an air of whispering, 'Come and find out'."*
>
> J. Conrad, *Heart of Darkness*, 1902

Experience

TOMRA Sorting NV — Haasrode, Belgium

Senior Research Scientist - PhD Student *Oct 2016 – Nov 2020*

Classical statistical methods (e.g. SVM's, PCA, QDA, CNN's, ...) often have poor performance on datasets with outliers. Although it is fairly easy to show the dramatic effects caused by just one outlier, the use of robust statistics in industrial machine learning tasks is only gradually being recognized in the last years. TOMRA Sorting develops industrial food inspection machines which process several gigabytes of high dimensional measurements in milliseconds, thereby pushing against the boundaries of the available computational power. My research activities already proved that the cross-fertilization of industrial machine classifiers and robust statistical methods was beneficial. However, the extreme high computational load continuously forces research into massive parallelization technologies and – more importantly - the research and development of new, robust and sparse machine learning methods which can process large datasets under real-time constraints.
To this end I initialised an intra disciplinary project between KULeuven, TOMRA and VLAIO which was directed towards strategic classification research with a clear defined economic finality. I was based at the MeBioS (Mechatronics, Biostatistics and Sensors) division, under the supervision of B. De Ketelaere, M. Hubert and P. J. Rousseeuw.

Keywords: Python/MATLAB, C++, Git, robust (non)linear regression and classification methods.

Senior Research Scientist *2014 – 2016*

GP-GPU's are powerful tools that are well-suited to unravelling complex, real-world problems, and form the foundation of today's deep learning frameworks such as TensorFlow & GPU accelerated optimization of SVM's. Due to the time criticality of our industrial classification tasks in day-to-day operations, combined with the vast amount of spectroscopic data, research into new, high performance (non-linear) modelling was essential. My research therefore focused specifically on the computational improvement of (ensembles of) LS-SVM's based on GPGPU's, which where then applied to spectroscopic datasets of various food related products. We effectively demonstrated the improved machine efficiency and computation times under real sorting conditions.

Keywords: MATLAB, C++, cuBLAS, CUDA, Git, global optimization, LS-SVM's, loss functions.

Research Scientist, digital signal processing *2010 – 2014*

Smart Sort is TOMRA's commercial name for an algorithm that I developed in two years. It enables machines to auto-configure itself: after providing training data the framework automatically ensures the most efficient machine settings for a given specificity-sensitivity trade-off setting. Although multi-class classification is entirely supported through ensemble learning, feedback from customers surprisingly learned us that it is mostly used as (an ensemble of) one class classifier(s). The algorithm is trained on good product(s) only, in turn demanding that the machine rejects all unknown, foreign materials such as glass, plastics and stones. It is currently still sold as a 6k€ machine option, and it became the de-facto classifier standard for tobacco sorting as it reduces machine change-over time dramatically. I implemented the first prototypes entirely in Java using test-driven development, and guided the software team towards it commercial/persistent implementation afterwards.

Keywords: Java, Octave, Subversion, optimization, regression, KDE, ensembles (boosting).

Unizo HQ — Brussels, Belgium

ICT-Consultant *2010 – 2010*

I managed ICT change processes for SME's related to the integration of ERP, CRM, webshop, .. software. I also took care of the daily incoming questions, as well as the organization of training courses, workshops and networking events.

Sint-Martinuscollege Overijse, campus Lipsius — Overijse, Belgium

Science and mathematics teacher — *2009 – 2010*

Par-time biology, maths - and science teacher for the 2' and 3' degree students.

Education

KU Leuven, faculty of Science — Leuven, Belgium

Doctor of Science (PhD): Statistics (dr.) — *2016 – 2020*

Development of real-time, robust statistical methods with novel applications in food sorting

KU Leuven, faculty of Engineering Science — Leuven, Belgium

Master of Electrical Engineering: Information Systems and signal Processing (ir.) — *2005–2009*

Development of an SVM-based OCR for Latin-Greek manuscripts

University of Antwerp, faculty of Biomedical Sciences — Antwerp, Belgium

Master of Biomedical Sciences: Neurosciences Research (lic.) — *2004 – 2005*

Reduction of ring artefacts on μ-CT images

Karel de Grote-Hogeschool — Antwerp, Belgium

Master in Electronics and ICT Engineering Technology (ing.) — *2001 – 2004*

Tristan: data acquisition software for a heat treatment production process

Karel de Grote-Hogeschool — Antwerp, Belgium

Bachelor of Electromechanics — *1998 – 2001*

Development of a (SQL-based) database application for the registration of production data

Publications

Published:

- I. Vranckx, De Ketelaere, B., M. Hubert, and P. J. Rousseeuw (2020). Development of real-time, robust statistical methods with novel applications in food sorting. Dissertation-thesis.
- I. Vranckx, J. Raymaekers, B. De Ketelaere, P. J. Rousseeuw, M. Hubert (2021). Real-time discriminant analysis in the presence of label and measurement noise. *Chemometrics and Intelligent Laboratory Systems 208*.
- De Ketelaere, B., M. Hubert, J. Raymaekers, P. J. Rousseeuw, and I. Vranckx (2020). Real-time outlier detection for large datasets by RT-DetMCD. *Chemometrics and Intelligent Laboratory Systems 199*.
- Raymaekers, J., P. J. Rousseeuw, and I. Vranckx (2018). Discussion of "The power of monitoring: how to make the most of a contaminated multivariate sample". *Statistical Methods & Applications 27*, 589–594.

Under review:

- J. Schreurs, I. Vranckx, B. De Ketelaere, M. Hubert, J. A.K. Suykens, P. J. Rousseeuw (2020). Outlier detection in non-elliptical data by kernel MRCD. *ArXiv e-prints*, arXiv:2008.02046. I. Vranckx and J. Schreurs contributed equally to this work.

Working papers:

- I. Vranckx, J. Schreurs. Robust and sparse support vector machines for data sets with label and measurement noise. I. Vranckx and J. Schreurs contributed equally to this work.

Skills & Interests (non-exhaustive)

Software interests/expertise: Recruitment of software and research engineers. Good knowledge of agile and Test Driven Development methodologies, functional and OO-based programming, version control. I enjoy writing Python/Java/Julia/C++ and Latex. Well experienced with industrial algorithm research, development & implementation. Sublime text, Visual Studio, Gitkraken and Pycharm user.

Natural languages: Dutch (*mother tongue*), English (*professional proficiency*), French (*elementary proficiency*)

Programming languages: MATLAB, Octave, R, Python, Julia, C++, CUDA, Java, SQL, R (RobustBase), ArrayFire, BLAS/LAPACK/LINPACK, Spark.

Research interests/expertise: Data mining, (robust) calibration methods, (non-linear) classification, robust statistics, big data, CNN's, real-time GP-GPU & MapReduce based processing, machine learning pattern recognition techniques, kernel embeddings of distributions, high dimensional, collinear data, Least Squares Support Vector Machines, anomaly & fraud detection, Kaggle datasets, multivariate global optimization methods, chemometrics, OOP-programming, HPC; External VLAIO expert.

General interests: Sourdough bread baking and jogging (i.e. my new COVID related hobbies), chess, BZFlag, Far Cry, cycling, typography (e.g. LaTeX), fraud detection, macrobiotic food, cooking, intra-disciplinary research collaboration, valorization of innovation, travel, speaking at conferences, lunch & learns, quality time with my three kinds.

Please refer to Github for updates of this document.