

“ 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’.”

Experience

J. Conrad, *Heart of Darkness*, 1902

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 sometimes 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 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.

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

Senior Research Scientist

2014 – 2016

GP GPU gebaseerde classificatoren. In essentie omvat zo een efficiëntie gedreven classifier een C++, GPU raamwerk waarin onderstaande technieken toegepast worden: (a) Detectie van multivariate outliers. (b) Normalisatie klasse balancerings – indien opportuun. (c) Ongesuperviseerde, statistisch aangedreven parameter optimalisatie. (d) Aansturen van het geïntegreerd crossvalidatiemodel. (e) Een Hammingafstand gebaseerde foutcorrectiemechanisme.

Keywords: MATLAB, C++, cuBLAS, CUDA, Python, Git, global optimization.

Research Scientist, digital signal processing

2012 – 2014

Research, Development en implementatie van het Smart Sort framework. Smart Sort is de commerciële naam van de wiskundige oplossingen die gebruikt worden om sorteermachines volledig automatisch in te stellen. Na het aanbieden van een trainingsdata zorgt dit raamwerk automatisch voor de meest efficiënte machine-instellingen. Zonder in detail te gaan omvat deze technologie: (a) Verschillende entropie, Hellinger - en mutual information gebaseerde contrastbepalingsfilters. (b) Verschillende outlier eliminatie principes om zuivere trainingsdata te krijgen. (c) Statistische efficiëntie gebaseerde (one, multiclass) classificatoren die specifieke normalisatiealgoritmes gebruiken. Deze technologie is integraal geïmplementeerd geworden in Java en communiceert rechtstreeks met de machinehardware over een generieke communicatielaag.

Keywords: JAVA, Octave, Subversion, optimization, regression, KDE

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

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

Please refer to my LinkedIn profile for a more complete list of work experience.

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 OCS for Latin-Greek manuscripts	
University of Antwerp, faculty of Biomedical Sciences	ANTWERP, BELGIUM
Master of Biomedical Sciences: Neurosciences (Research)	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. Development of real-time, robust statistical methods with novel applications in food sorting (2020). 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.
-

Skills

Technical expertise: Leading and recruiting teams of software engineers. Big fan of Agile methodologies, continuous delivery and functional programming. Enjoys writing Ruby/Python/Java/C++ and Haskell. Solid knowledge of the full web technology stack. Able to architect *and* implement distributed/HA systems. Strong Linux administration skills (e.g. Bash scripting, Apache/NGINX, Postgres/My/NoSQL, Elasticsearch). Well experienced with virtualization/containerization (Docker/Kubernetes, KVM, Xen and several AWS solutions) and DevOps (Puppet). Sublime text, Gitkraken and Pycharm user.

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

Interests

Research interests (non-exhaustive): 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.

General interests (non-exhaustive): Sourdough bread baking (COVID related hobby), chess, OOP-programming, jogging, cycling, typography (e.g. \LaTeX), fraud detection, macrobiotic food, cooking, intra-disciplinary research collaboration, valorization of innovation, travel, speaking at conferences.