

“ 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 (VLAIO Baekeland mandaat)

Oct '16 – present

Classical machine learning methods (e.g. SVM's, PLS, PCA, 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 in the majority of statistical machine learning techniques, the use of robust statistics in industrial machine learning tasks is only gradually being recognized in the last couple of years. Modern industrial food sorting machines are processing several gigabytes of high dimensional chemometrical measurements in milliseconds, pushing against the boundaries of the available computational power. Furthermore, the amount of data is expected to grow over the coming years. The cross-fertilization of industrial machine classifiers and robust statistical methods in a big data scene has undoubtedly proven to be beneficial. However: the high computational demand forces the continuous exploration of new massive parallelization technologies and – more importantly - the research and development of new, robust sparse statistical methods which can handle modern datasets under state of the art, industrial real-time constraints.

Programming languages: Python, MATLAB, C++, Julia, Git

Senior Research Scientist

Oct '16 – present

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

Programming languages: MATLAB, C++, cuBLAS, CUDA, Python, Git

Research Scientist, digital signal processing

Oct '16 – present

Chemometrische analyse, onderzoek en implementatie van chemometrische modellen, pre-processing. Mijn initieel werk was het onderzoek naar de werking en haalbaarheid van chemometrische methoden (PCR, O-PLS, PLS-DA, ...) en pre processing methoden die goede resultaten gaf op onze 256-dimensionale spectrale data (er bestaan immers geen specifieke stel-regels die onmiddellijk leiden tot het vinden van de beste methode). Prototypes werden door mij geïmplementeerd in Java. Daarnaast had ik een groot aandeel in de inverse engineering van bestaande classificatoren. 2. 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 efficiënte gebaseerde (one, multiclass) classificatoren die specifieke normalisatie algoritmes gebruiken. Deze technologie is integraal geïmplementeerd geworden in Java en communiceert rechtstreeks met de machine hardware over een generieke communicatielaag.

Programming languages: JAVA, Octave, batch scripts, Subversion

Unizo HQ

BRUSSELS, BELGIUM

ICT-Consultant

2010 – 2010

Als IT adviseur stond ik in voor de software en hardwarematige ICT-advisering aan KMO's op basis van input van externe ICT-partners, eigen onderzoekwerk en eigen ervaringen. Ik initieerde, begeleidde en bracht ICT-veranderingstrajecten (ERP, CRM, CMS, ...) tot uitvoering bij KMO's en ruimer binnen een sector. Naast de bovengenoemde punten verzorgde ik de dagelijkse communicatie en opvolging met de betrokken beroepssectoren - evenals de organisatie van zowel praktische als inhoudelijke opleidingen, workshops en netwerkevents

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.

General interests (non-exhaustive): art, Buddhism, cryptography, functional programming, Go (board game), history, music (from classical and jazz to Berlin-techno), NLP, permaculture, philosophy, rock climbing, startups, travel, typography (e.g. graphic design, \LaTeX), UX-design and vegan cuisine.