efl insights 01|2021
An efl – the Data Science Institute Publication

efl | the Data Science Institute

Start Thinking in Systems

Buying into Fraud – German Retail Investors and the Wirecard Scandal

Insights from Explainable Interactive Machine Learning in the Age of COVID-19

The Customer Determines the Success or Failure of the Company

DEUTSCHE BÖRSE GROUP
DZ BANK Gruppe
finanz informatik
AMERICAN EXPRESS
Deutsche Leasing
FACTSET

GOETHE UNIVERSITÄT FRANKFURT AM MAIN
TECHNISCHE UNIVERSITÄT DARMSTADT

Impressum

Redaktion
Prof. Dr. Peter Gomber
Dr. Jascha-Alexander Koch

Herausgeber
Prof. Dr. Wolfgang König
Vorstandsvorsitzender
efl – the Data Science Institute e. V.

Prof. Dr. Peter Gomber
Stellvertretender Vorstandsvorsitzender
efl – the Data Science Institute e. V.

Kontakt
info@eflab.de
www.eflab.de

Gestaltung
Novensis Communication GmbH
Bad Homburg

1. Ausgabe, 2021
Auflage: 200 Stück (Print) / 2.000 Stück (E-Paper)

Copyright © by efl – the Data Science Institute e. V.

Printed in Germany
ISSN 1866-1238

Editorial

Start Thinking in Systems

Berthold Kracke

Dr. Berthold Kracke
Head of Global Operations
Clearstream

The financial industry is in the process of transformative change. Driven by changing customer needs, cost pressure, and regulatory change, this decade is about to bring a technological transformation that will lead us from an old world, characterized by redundancy, fragmentation, and reconciliation requirements, to a new world of seamless interaction between investors and issuers.

This transformation will be based on a number of technologies whose names have become so familiar that I hesitate to call them "new": artificial intelligence, machine learning, cloud, distributed ledger technology, blockchain. These tools will allow us not only to digitize our systems and processes, but ultimately also our products.

Being responsible for global operations at Clearstream, a leading provider of post-trade market infrastructure, one area of focus for me is the continuous improvement of our processes. Traditionally, our industry has been very operations-heavy and grapples with a lot of manual processes which take up resources and are vulnerable to human error. As cost pressure mounts, market participants are increasingly looking to new technologies for their potential to create operational efficiencies.

How do we make that work in practice? If we want the promise of automation and machine learning to materialize on a large scale, it is vital for us not to see individual use cases or technologies in a vacuum. While it is important to focus on concrete use cases that can bring real benefit, that alone is not enough. We need to create the right ecosystem for these solutions to work.

Achieving true technological transformation in post-trading operations is conditional on providing the right expertise, the right technical infrastructure, and the right contractual framework. To be successful, you need to employ experienced data scientists and engineers who will be able to support your business in developing machine learning and automation solutions. Provide access to flexible cloud services as the necessary technical foundation. Ensure that these services are available within a highly regulated environment with strict data protection requirements.

At Clearstream, we have been able to successfully implement machine learning solutions in production that have resulted in tangible efficiency gains, e.g., by using machine learning in SWIFT message routing to improve straight-through processing rates or by employing optical character recognition and natural language processing to classify and process incoming e-mail attachments in fund operations.

We have been able to do so because, being a part of Deutsche Börse Group, we are embedded in an excellent ecosystem with a clear focus on new technologies, which provides centralized big data and automation tools and services as well as agile development with mixed business and IT teams. This is further supported by Deutsche Börse's multi-cloud strategy, which enables us to work in a regulated environment while making use of cloud services provided by trusted partners.

Harnessing the potential of automation and machine learning takes a holistic view that considers the interaction of technologies and the necessary legal and regulatory framework – ensuring a focus on stability and integrity at least as much as on efficiency. This holds true for all aspects of technological transformation. It is a challenge we can only take on if we stop seeing processes and technologies in a vacuum and instead start thinking in systems.

Research Report

Buying into Fraud – German Retail Investors and the Wirecard Scandal

ON JUNE 18TH, WIRECARD'S SHARE PRICE PLUMMETED BY MORE THAN 60% FOLLOWING THE FIRM'S ADMISSION OF BEING SUBJECT TO "ENORMOUS FRAUD" AND BILLIONS OF EUROS MISSING. THIS REPORT DOCUMENTS GERMAN RETAIL INVESTORS' RESPONSE AND FINDS THAT THE POPULARITY OF WIRECARD AMONG RETAIL INVESTORS LED TO SUBSTANTIAL LOSSES IN THEIR PORTFOLIOS. THESE LOSSES WERE EXACERBATED BY STRONG BUYING SENTIMENT AFTER THE ANNOUNCEMENT. THE FAILING STOCK WAS PURCHASED BY INVESTORS ALREADY ENGAGED IN IT AS WELL AS NON-EXPOSED CUSTOMERS.

Konstantin Bräuer

Andreas Hackethal

Guido Lenz

Thomas Pauls

Introduction

In an unprecedented event, Wirecard, a member of the prestigious German DAX index, lost more than 60% of its market value in a single day and became virtually worthless in subsequent trading days. This day, June 18th, was the first time Wirecard's board admitted missing escrow accounts worth billions of Euros and resembles the finale in a long story of accusations and denials of fraud. This report dissects the investment reaction of German retail investors on the largest corporate scandal in recent history destroying billions in market capitalization in a matter of days.

Long History of Allegations

First allegations of balance sheet irregularities concerning the payment service company surfaced already in 2008 from a German shareholder association. In the following years, Wirecard grew fast and expanded internationally fueled by acquisitions mainly in Asia. Yet, the questions raised on irregularities in the books became more evident. Most prominently, the Financial Times started to publish its series "House of Wirecard" in 2015 followed by a report from the short investor "Zatarra" in 2016. However, the stock price was never impressed by these allegations and, supported by Wirecard's denials, rose to a peak of EUR 195 in August 2018 valuing the company close to EUR 25 billion. Only a month later, Wirecard replaced Commerzbank in the German DAX index and was finally considered the most important and successful German (Fin-)Tech startup. Despite BaFin banning short sales on Wirecard for two months in 2019, the pressure of public fraud allegations increased, and the stock price declined to around EUR 100 at the end of the year. Reports for both audits, the annual 2019 and a special one to investigate the allegations on cash positions, were delayed multiple times and finally not exonerating the company (DGAP, 2020). After the police searching Wirecard's headquarters, the management released a video statement addressing missing escrow accounts, an "enormous fraud case", and the failure to get the balance sheet certified on June 18th. The same day, the COO was suspended and the following day, the CEO stepped down being arrested shortly after. This was the first admission of irregularities by the company and led the firm's value to vanish within days leading to the filing of bankruptcy only a week later on June 25th.

Data

We obtain data from a German bank with a holistic offering in retail banking services. The bank provides us with data including customer demographics and administrative records of individual customer's financial assets, portfolio holdings, and security transactions. The complete sample period ranges from July 2017 until July 2020 and includes a five-digit number of individual investors of which roughly 5% directly held shares of Wirecard on June 17th, the day before the admission of fraud and the price drop.

Who Was Invested Beforehand?

Examining the cross-sectional characteristics of investors directly holding Wirecard shares on June 17th, right before the release of the board's statement, we find significant differences compared to non-holders. Wirecard holders are more likely to be male and tend to be more affluent in terms of their financial assets managed by the bank. In fact, their average net wealth is more than double the amount of the non-holders, which is driven by both a 135% higher value within their brokerage accounts and a 76% larger amount of deposits. Furthermore, they also have an almost 20% higher regular net income. In line with holding a position in a single stock, the self-reported risk aversion of Wirecard investors is lower, i.e., they consider themselves more tolerant towards risk, than investors not holding Wirecard shares. This higher risk tolerance translates into a 35% higher propensity to hold single stocks and a 50% higher share of single stocks in their overall portfolios. Consequently, Wirecard investors also show on average five times higher monthly activity in their brokerage accounts resulting in an average of two monthly trades. A further indicator of their experience in investment decisions can be derived from their likelihood of investing into passive investment vehicles which is almost three times higher than for non-Wirecard investors. These differences in investor and portfolio characteristics point towards a selection of wealthier, less diversified individuals, who are more experienced in capital markets and allo-

cate deliberately a larger amount into Wirecard as a single investment. This single position comprises on average EUR 17,000 or 7% of their portfolio values. Relative to the book-value of their holdings on June 17th, Wirecard investors lost on average more than EUR 16,000 until the end of June. The median loss of EUR 5,000 and losses exceeding EUR 34,000 for 10% of all investors clearly show a highly right skewed distribution of losses. On the other side, only very few investors could limit their losses by selling quickly during the fast drop in stock price. And even more investors enlarged their losses by increasing their Wirecard exposure on June 18th.

How Did Investors Respond to the Share Price Drop?

Figure 1 shows daily transaction volumes of our investor sample in Wirecard shares. Strikingly,

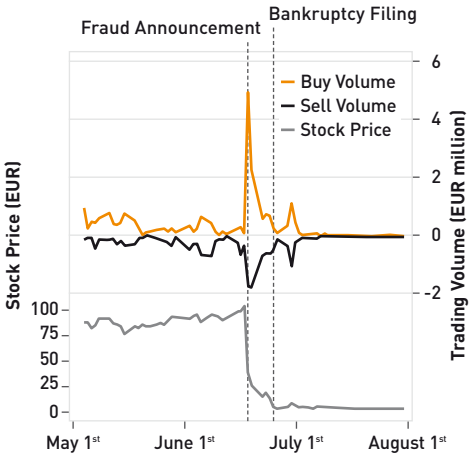

Figure 1: Daily Stock Price and Trading Volume of Wirecard over Time

on June 18th, buying volumes are almost three times as high as those of sales despite the board's announcement. On this very day, almost 700 investors bought Wirecard shares worth roughly EUR 5 million while sales only summed up to less than EUR 2 million (by 189 investors). The average transaction volume was around EUR 7,800. This substantial shift in the buy-sell-imbalance is present in Euro volumes, number of investors trading, as well as transactions, and implies a strong positive sentiment amongst retail investors despite the public announcement of the firm being part of an "enormous fraud". Most of the sampled retail investors apparently assigned only a very small probability to a potential aggravation of Wirecard's position and instead used the chance to (re)purchase at lower prices. The significant increase in purchasing volume was driven not only by existing investors (i.e., those who held shares as of June 17th), but also by new investors purchasing Wirecard for the first time. This increased the number of Wirecard holders in the sample by more than 10% as shown in Figure 2. In this figure, it is also evident that a large "flight" out of Wirecard only occurs a week later when the bankruptcy filing was announced and more than a third of investors liquidated their positions entirely. These results reflect recent evidence on retail investors' trading behavior during the COVID-19-induced stock market crash in which retail investors "acted as a (small) market-stabilizing force" (Welch, 2020). Despite the settings being vastly different, our sampled retail investors also provided a certain liquidity during times of high uncertainty to the market. This is

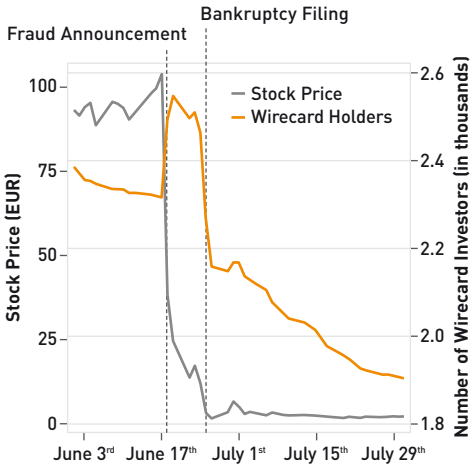

Figure 2: Daily Stock Price and Number of Investors Holding Wirecard Shares

quite remarkable since Kuvvet (2015) finds that market liquidity is depressed following corporate fraud events.

Conlusion

With their investment in Wirecard, many retail investors experienced substantial losses. Since we generally find that these affected investors are on average more experienced, more affluent, and less risk-averse than non-invested customers, their relative loss remains, despite total losses in Wirecard for many, seemingly manageable. This indicates that these investors are (at least partially) aware of downside potentials in single stock positions. As a next step, we investigate the impact of this negative wealth shock on investor behavior in both the financial realm as well as their everyday life. Interesting questions to answer on subsequent financial decision-making include, for example: Do affected investors trade less after the shock or reduce their exposure to risky assets further? Do they change their asset allocation or security selection towards less idiosyncratic risk? For affected investors' everyday life, it will be interesting to investigate whether these losses translate into adjustments of consumption and/or savings behavior as well as consequences in the standard of living. On the aggregate market level, it will be interesting to see whether retail investors can provide market stabilizing liquidity in uncertain times.

References

Kuvvet, E.:
Corporate Fraud and Liquidity.
In: The Journal of Trading, 10 (2015) 3, pp. 65-71.

Welch, I.:
Retail Raw: Wisdom of the Robinhood Crowd and the COVID Crisis.
In: NBER Working Paper, w27866, 2020.

Financial Times:
The House of Wirecard.
https://ftalphaville.ft.com/2015/04/27/2127427/the-house-of-wirecard/, 2015.

DGAP:
Wirecard AG: Wirecard AG verschiebt Jahresabschluss 2019.
https://www.dgap.de/dgap/News/corporate/wirecard-wirecard-verschiebt-jahresabschluss/?newsID=1353571, 2020.

Research Report

Insights from Explainable Interactive Machine Learning in the Age of COVID-19

COVID-19 HAS AGAIN TIGHTENED ITS GRIP AROUND THE WORLD AND THE HEALTH SYSTEM. THIS ARTICLE GIVES AN INTRODUCTION TO EXPLAINABLE INTERACTIVE MACHINE LEARNING AND PROVIDES INSIGHTS ON HOW THIS METHOD MAY NOT ONLY HELP IN ENGINEERING MORE POWERFUL AI SYSTEMS, BUT ALSO HOW IT MAY HELP TO EASE THE BURDEN OF VIRAL STRAINS ON THE HEALTHCARE SYSTEM.

Oliver Hinz
Wolfgang Stammer
Benjamin M. Abdel-Karim
Christian Hügel
Kristian Kersting
Nicolas Pfeuffer
Patrick Schramowski
Andreas Bucher
Gernot Rohde

Introduction

The latest reports of the World Health Organization on the coronavirus pandemic (www.who.int) show alarming numbers of millions of new infections every week. Although these numbers may reflect the severity of the outbreak, they do not yet reveal the dimensions of the burden the virus has on the health system. Recent reports from the German center of intensive healthcare (www.intensivregister.de/#/reporting) show that, by November 2020, the amount of required beds and intensive care unit slots for COVID-19 patients has spiked from a low between May and October to unseen high numbers. These numbers are just an indication of how much healthcare workers on the forefront are exceedingly fighting for the lives of the patients in an ever-growing fear of an unmanageable situation (e.g., Lai et al., 2020).

Machine Learning as a Supportive Pillar for Healthcare

To support our healthcare system and workers in the fight against the pandemic, scholars from the field of computer science, information systems, and medicine continuously try to engineer machine learning (ML)-based clinical decision support systems (CDSS). In the course of the year 2020, a wealth of research projects and papers has been published, ranging from aggregated open datasets to support the development of CDSS against COVID-19, implementation of black-box (e.g., Chowdhury et al., 2020) and white-box approaches, as well as experiments on the usefulness of such systems versus humans and in human-machine hybrid constellations (e.g., Mei et al., 2020).

The pitfalls of ML-Based CDSS

Arguably, many studies have reported promising results of either novel or existing architectures, e.g., convolutional neural networks (CNN) in the detection of COVID-19 from X-rays (e.g., Chowdhury et al., 2020) or CT-Scans (e.g., Mei et al., 2020), or for human-hybrid constellations (e.g., Mei et al., 2020). Nevertheless, two evident problems of these CDSS solutions are (1) that physicians cannot correct or improve these systems based on their expertise, and (2) that many of these proposed systems come with a lack of transparency.

As a result, these systems are not only potentially overconfident, but also dangerously intransparent. Hence, the employment of such systems in time-pressing, stressful environments could make up a potentially dangerous combination.

Our Methodology: XIL to the Rescue

Against this background, we aim to remedy the shortcomings of current systems with the aid of a methodology called explainable interactive machine learning (short: XIL, see Figure 1).

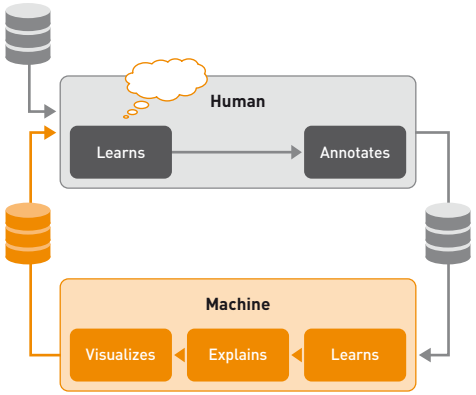

Figure 1: XIL Cycle (adapted from Hinz et al., 2020)

Grounded on the concepts of explainability (e.g., Selvaraju et al., 2017) and interactive machine learning, recent experiments with biologists had shown the potential of XIL in not only making black box systems, such as CNN, more transparent, but also its potential in helping experts to make the system more accurate and satisfying to use (Schramowski et al., 2020). Apart from the advantages in engineering, by augmenting ML systems with explainable AI features (xAI) and putting the human in the loop, XIL also has the potential to help humans discover and learn novel facts about the underlying task. Since knowledge about COVID-19 is still incomplete, XIL may serve as a valuable method for exploring and gaining novel knowledge. Our methodology consisted of several stages: first, we

selected a suitable database for engineering a helpful CDSS for the case of COVID-19-based pneumonia.

In this matter, we oriented ourselves according to the diagnostic guidelines of the British society of thoracic imaging (Nair et al., 2020), and thus decided to build a CNN with the target of predicting COVID-19-based pneumonia from X-rays. Our selected database (Chowdhury et al., 2020) consisted of three classes: normal (X-rays without clear signs of pneumonia), viral pneumonia (pneumonia caused by other viral pathogens), and COVID-19-based pneumonia (pneumonia caused by the SARS-CoV-2 pathogen). As a foundational architecture, we used the ImageNet pre-trained AlexNet architecture (Krizhevsky et al., 2017). With initial fine-tuning of the classifier on our dataset, we, then, engage in several cycles of XIL to improve the classifier as well as to try to generate novel insights.

Let the XIL Cycle Begin

The XIL process has the advantage that, after each training cycle, the results and the corresponding explanations for classification can be inspected and can, then, be readjusted according to the human's expertise. Our team consisted of computer scientists, information systems researchers, as well as radiologists and pneumologists. In the case of our application area, namely the detection of COVID-19-based viral pneumonia, especially pneumologists and radiologists played a vital role in assessing potential errors that could bias the system and ultimately put its usefulness at risk.

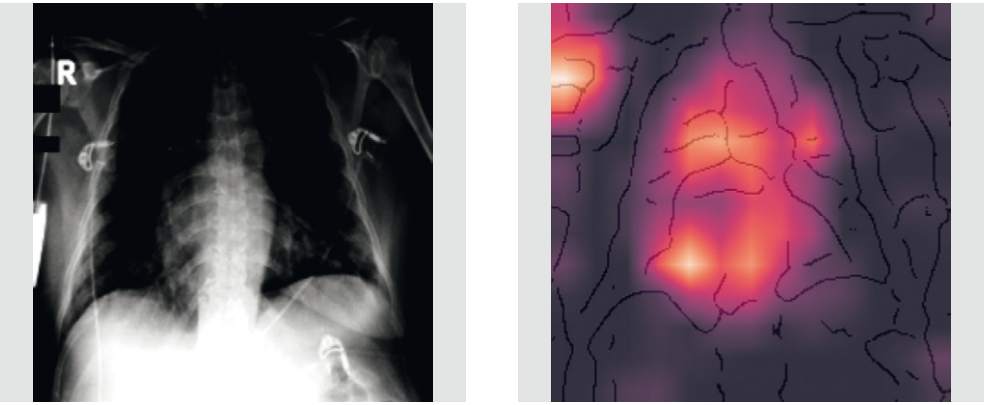

Figure 2: Original X-Ray on the Left, X-Ray with Grad-CAM Overlay on the Right (highlighted regions indicate a focus of the CNN on these salient areas)

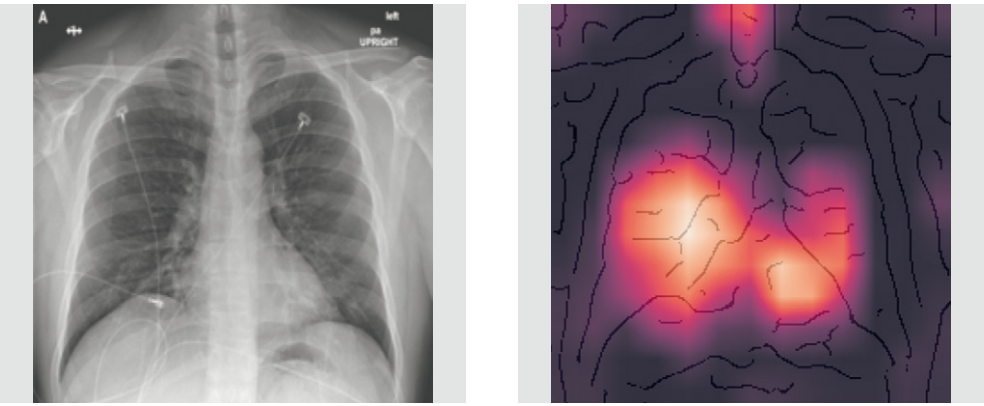

Figure 3: Original X-Ray on the Left, X-Ray with Grad-CAM Overlay on the Right (highlighted regions show the effect of annotations and penalization)

Cycle 1

At the beginning of Cycle 1, we train the CNN with the available X-ray images. Our resulting classifier showed high performance metrics, thus, indicating a very good classifier (accuracy: 91.06%, precision: 93.64%, recall: 92.53%). With the aid of explainability features, we are, then, able to provide explanations for each classified image in the test set. For generating the explainability, we applied the Captum library (see https://captum.ai/). We initially used Grad-CAM (e.g., Selvaraju et al., 2017), providing coarse results (see Figure 2), but, then, switched to Captum's VarGrad implementation to which we added a self-developed red/blue filter to provide a more precise explanation (e.g., Figure 4).

After thorough inspection of the test cases by the physicians and the rest of the team, we concluded that several obvious confounders (such as letters or medical instruments in the images, see the "R" in Figure 2) needed to be excluded or penalized in the model.

Cycle 2

We, thus, proceeded to refine the database of the model by letting crowdworkers from Mechanical Turk annotate important regions for the classification (torso, lungs, throat) in about 3,000 X-rays. In combination with the generated annotations, we fine-tuned the CNN again – now penalizing the classifier for unimportant regions (Figure 3 shows now that the algorithm focuses on the actual areas of interest). We again arrived at a very good model as indicated by the performance metrics (accuracy: 94.50%, precision: 96.30%, recall: 92.93%). After thorough inspection with the help of explanations in Cycle 2, the research team discovered another important issue: apparently, the used data were systematically biased. The labelled classes for the 'normal' and 'viral pneumonia' consisted of mainly pediatric images, while the 'COVID-19' class consisted of people of mixed age. This issue only became apparent, since the explanation features indicated that the algorithm looked at specific

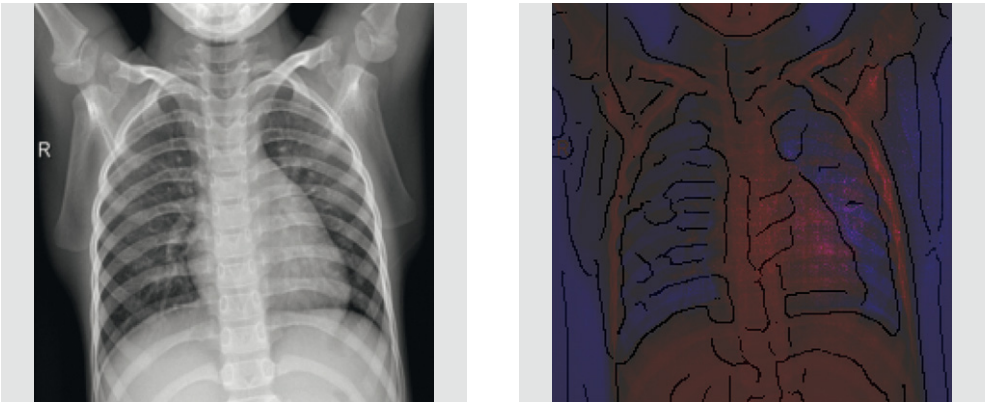

Figure 4: Original X-Ray on the Left, X-Ray with VarGrad Red/Blue Saliency Map Overlay on the Right (highlighted regions indicate a focus of the CNN on these salient areas)

skeletal characteristics of children; notably, we discovered that this issue, thus, persists in various published papers that used the same database but they were not able to recognize this potentially critical flaw.

Cycles 3a and 3b

The findings of XIL Cycle 2 lead us to subsequently revise the database in different ways, discarding the 'COVID-19' class in Cycle 3a to inspect explanations and accuracy of a binary classifier for viral pneumonia, as well as shifting to a different database in Cycle 3b (Wang et al., 2017). In doing so, we achieve a less accurate, but yet potentially more generalizable and meaningful model in this final cycle.

Summary of Results

To summarize our research efforts, we explored the utility of employing XIL in the construction of a mature CDSS for the case of classifying COVID-19-based viral pneumonia from X-rays. In our research efforts, we are able to show the benefits of XIL, namely a deep inspection of the classifier, not only providing accountability and transparency of the classifier, but also helping to uncover potentially dangerous confounders, which helps to finally arrive at realistic and promising models for practice. In a general perspective, our insights are not only important for the medical domain or for image data, but they foreshadow the potential of a promising novel methodology for every kind of data and domain. For example, XIL may also be highly helpful for the assessment of different kinds of algorithmic bias in ML-based systems to correct systems and may, thus, help organizations in fulfilling regulatory demands. In conclusion, we are confident that XIL may be further explored by science and adopted in practice to contribute to the usage of accurate, unbiased, and transparent ML-based systems.

References

Chowdhury, M. E.; Rahman, T.; Khandakar, A.; Mazhar, R.; Kadir, M. A.; et al.:
Can AI Help in Screening Viral and COVID-19 Pneumonia?
In: IEEE Access, 8 (2020), pp. 132665-132676.

Hinz, O.; Pfeuffer, N.; Stammer, W.; Schramowski, P.; Abdel-Karim, B. M.; et al.:
How Explainable Interactive Machine Learning Can Solve the Most Pressing Problems in Machine Learning – The Example of Diagnosing Viral Pneumonia.
In: Working Paper, 2020.

Krizhevsky, A.; Sutskever, I.; Hinton, G. E.:
ImageNet Classification with Deep Convolutional Neural Networks.
In: Communications of the ACM, 60 (2017) 6, pp. 84-90.

Lai, J.; Ma, S.; Wang, Y.; Cai, Z.; Hu, J.; et al.:
Factors Associated with Mental Health Outcomes Among Health Care Workers Exposed to Coronavirus Disease 2019.
In: JAMA Network Open, 3 (2020) 3, e203976, pp 1-12.

Mei, X.; Lee, H. C.; Diao, K. Y.; Huang, M.; Lin, B.; et al.:
Artificial Intelligence-Enabled Rapid Diagnosis of Patients with COVID-19.
In: Nature Medicine, 26 (2020) 8, pp. 1224-1228.

Nair, A.; Rodrigues, J. C. L.; Hare, S.; Edey, A.; Devaraj, A.; et al.:
A British Society of Thoracic Imaging Statement: Considerations in Designing Local Imaging Diagnostic Algorithms for the COVID-19 Pandemic.
In: Clinical Radiology, 75 (2020) 5, pp. 329-334.

Schramowski, P.; Stammer, W.; Teso, S.; Brugger, A.; Herbert, F.; et al.:
Making Deep Neural Networks Right for the Right Scientific Reasons by Interacting with Their Explanations.
In: Nature Machine Intelligence, 2 (2020) 8, pp. 476-486.

Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D.:
Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization.
In: Proceedings of the IEEE International Conference on Computer Vision (ICCV); Venice, Italy, 2017, pp. 618-626.

Wang, X.; Peng, Y.; Lu, L.; Lu, Z.; Bagheri, M.; Summers, R. M.:
ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases.
In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Honululu (HI), US, 2017, pp. 2097-2106.

Insideview

The Customer Determines the Success or Failure of the Company

INTERVIEW WITH PHILIPP SCHMITT

Dr. Philipp Schmitt
Vice President
American Express Europe S.A.
(Germany Branch)

Companies in the financial industry are under enormous pressure to innovate. What does a company have to do to keep up with the pace?

In addition to the immediate response to the COVID-19 pandemic, the financial services industry is currently facing three fundamental challenges: increased regulation, the digitization of all areas of life, and permanently low interest rates. It is essential to think from the customers' perspective and find new ways to use technology to make life easier for them. And ideally, to generate more revenue and/or to reduce costs.

What are the customers' needs today?

They want to pay faster, more seamlessly, and safer. From a customer's point of view, frictionless payments are great. One-click checkouts at online shops, voice shopping, and self-checkouts in supermarkets have made the payment process a new experience. Contactless payment has increased significantly during the pandemic. It is easy and comfortable. Probably one day we will be able to pay with the finger or face. These solutions help to make everyday life easier.

The new payment methods are simple. Are they secure?

Among other things, we rely on artificial intelligence (AI) to protect our customers from fraud. Transaction data is analyzed in real time to gain insights that humans would miss. An AI-based system uses these insights to improve itself. An example: today, a delivery of ski boots to the Maldives should not be a problem for fraud systems. AI analyzes customer behavior quite precisely. For example, we look at the pauses between keystrokes and try to reliably recognize our real customer. AI can help to identify the real customer among the conspicuous transactions, approve the correct transactions, and perform this in the fractions of a second that payment transactions take today.

Electronic payments generate incredible data insights. Is this a risk to data protection?

Transparency does not mean sharing data just because it is possible. We have the responsibility to make consumers aware of both the opportunities and the risks. Data privacy protection laws are stricter in Germany than in most other countries. However, 51% of consumers are not making use of certain online offers because they are worried about data security. We need to be aware how valuable this data is. Trust is the key. To whom do I give access to my data as a customer? Trustworthy brands will succeed.

Are there other demands customers require from the financial service industry?

Payments must be more than just moving money from point A to point B. It is also about additional benefits. Merchants want more service, companies want support for supplier payments, and consumers want unique experiences.

Where do all these changes lead to?

Firstly, more collaboration. Innovation happens through both competition and collaboration. We see established players, FinTechs, and companies outside the financial service industry collaborating. Secondly, the customer will become even more front and center of all decisions. The customer determines the success or failure of the company. This insight is fortunately becoming more prevalent in the corporate world, accelerated by increasing choice and nearly complete transparency of options. The implications are significant for all areas: financial reporting, organizational structure, corporate strategy to name just a few. Only those who put the customer first in everything will succeed.

Thank you for this interesting conversation.

Infopool

News

efl Annual Conference 2021 Postponed Due to COVID-19 Pandemic
Our annual conference will not take place in spring 2021 as scheduled. The board of the efl – the Data Science Institute has decided to postpone the efl Annual Spring Conference due to the COVID-19 pandemic. We will update you on the potential timing and format of the conference in our next newsletter. Please, stay healthy!

efl Alumnus Marten Risius Receives AIS Early Career Award 2020
Dr. Marten Risius, Senior Lecturer at University of Queensland, received the 2020 AIS Early Career Award. Marten received the Information Systems discipline's highest distinction for early career academics only four years after receiving his Ph.D. from the efl – the Data Science Institute, despite a 7-year eligibility period. Congratulations!

Daniel Blaseg Receives Dissertation Prizes
Dr. Daniel Blaseg received the "Wolfgang Ritter Prize 2020" and the "Roman Herzog Forschungspreis Soziale Marktwirtschaft" for his dissertation "Crowdfunding". Congratulations!

Two Best Paper Awards
The paper "Mutual Funds and Risk Disclosure: Information Content of Fund Prospectuses" co-authored by Jonathan Krakow (University of Zurich) and Timo Schäfer (efl) received this year's Swiss Finance Institute Best Paper Award. Hendrik Jöntgen wins the Best Paper Award at the European Conference on Information Systems 2020 with his article "Fueling the Firestorm: Effects of Social Capital on Users' Persuasiveness during Online Firestorms".

Successful Disputation
Gabriela Alves Werb successfully defended her dissertation on "Measuring Risks for Consumers, Firms, and Investors in the Digital Economy" at the interface between information systems, finance and marketing.

Four New Colleagues
Jonas De Paolis joins the Chair of Prof. Gomber as a research assistant in January 2021. His research interests lie in market microstructure and complex financial networks. Emily Kormanyos, Fabian Nemeczek and Jan Radermacher joined the Chair of Prof. Hackethal as research assistants in summer 2020. Emily's interests lie in the cross-section of econometrics and data science. Fabian will focus on information retrieval and processing of individual investors as well as on investment flows, and Jan's research focus is the application of machine learning techniques to household finance.

Data Science and Pandemic Management
Prof. Hinz (together with Prof. Rhode, Prof. Drosten, Prof. Ciesek and others) joins the "Nationales Forschungsnetzwerk der Universitätsmedizin" as PI for digitization and data science in the project EViPan for the development, testing and implementation of regionally adaptive care structures and processes for evidence-based pandemic management.

Selected efl Publications

Abdel-Karim, B. M.; Pfeuffer, N.; Hinz, O.:
Machine Learning in Information Systems Research – A Systematic Review and Open Research.
Forthcoming in: Electronic Markets, 2021.

Abdel-Karim, B. M.; Pfeuffer, N.; Rohde G.; Hinz, O.:
How and What Can Humans Learn from Being in the Loop? Invoking Contradiction Learning as a Measure to Make Humans Smarter.
In: German Journal on Artificial Intelligence, 34 (2020) 2, pp. 199-207.

Bang, T.; May, N.; Petrov, I.; Binnig, C.:
AnyDB: An Architecture-less DBMS for Any Workload.
Forthcoming in: Proceedings of the 11th Annual Conference on Innovative Data Systems Research (CIDR 21); virtual event, 2021.

Clapham, B.; Gomber, P.; Lausen, J.; Panz, S.:
Liquidity Provider Incentives in Fragmented Securities Markets.
In: Journal of Empirical Finance, 60 (2021), pp. 16-38.

Han, S.; Reinartz, W.; Skiera, B.:
Capturing Brand and Customer Focus of Retailers.
Forthcoming in: Journal of Retailing, 2021.

Hanspal, T.; Weber, A.; Wohlfart, J.:
Exposure to the COVID-19 Stock Market Crash and its Effect on Household Expectations.
Forthcoming in: Review of Economics and Statistics, 2021.

Jöntgen, H.:
Fueling the Firestorm: Effects of Social Capital on Users' Persuasiveness during Online Firestorms.
In: Proceedings of the 28th European Conference on Information Systems (ECIS); Marrakech, Morocco, 2020.

Jasny, M.; Ziegler, T.; Kraska, T.; Röhm, U.; Binnig, C.:
DB4ML – An In-Memory Database Kernel with Machine Learning Support.
In: Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data; Portland (OR), US, 2020, pp. 159-173.

Krakow, J.; Schäfer, T.:
Mutual Funds and Risk Disclosure: Information Content of Fund Prospectuses.
In: 47th Annual Meeting of the European Finance Association (EFA); Helsinki, Finland, 2020.

Teso, S.; Hinz, O.:
Challenges in Interactive Machine Learning – Toward Combining Learning, Teaching, and Understanding.
In: German Journal on Artificial Intelligence, 34 (2020) 2, pp. 127-130.

van der Aalst, W.; Hinz, O.; Weinhardt, C.:
Big Digital Platforms.
In: Business & Information Systems Engineering, 61 (2019) 6, pp. 645-648.

For a comprehensive list of all efl news postings see: *https://www.eflab.de/news*
For a comprehensive list of all efl publications see *http://www.eflab.de/publications*

Infopool

RESEARCH PAPER: DEEP LEAKAGE FROM GRADIENTS

Federated Machine Learning is a new approach that aims to protect the privacy of data owners in a collaborative machine learning setup since only model gradients, and not training data, need to be revealed to others. However, this paper shows that it is possible to construct a privacy attack that allows a central parameter server, that is typically used in federated learning, to reconstruct individual training examples (e.g., a training image of one participant). This attack only requires access to the gradients that clients need to reveal, and to the global machine learning model that is anyway known to all participants. The authors show that the attack is applicable to different types of data and briefly review various defense strategies.

Zhu, L.; Liu, Z.; Han, S.
In: Advances in Neural Information Processing Systems 32, Proceedings of the Annual Conference on Neural Information Processing Systems 2019; Vancouver, BC, Canada, 2019.

RESEARCH PAPER: THE INFLUENCE OF PROFESSIONAL SUBCULTURE ON INFORMATION SECURITY POLICY VIOLATIONS – A FIELD STUDY IN A HEALTHCARE CONTEXT

Considering that most data breaches are ascribable to human action, researchers are keen to understand why individuals abide by or violate information security policies (ISP). This article investigates different attitudes toward ISP violations among three prominent professional healthcare groups: physicians, nurses, and support staff. The authors find substantial differences in intentions and behaviors regarding ISP violations based on professional subcultures. One such behavior is pseudocompliance, which appears to be compliant with ISP on the surface, but is an intentional ISP violation. A classic example is the behavior of employees to create a password that technically complies with a password ISP (e.g., using a complex password) but, then, write down the password and store it in an insecure place.

Sarkar, S.; Vance, A.; Ramesh, B.; Demestihas, M.; Wu, D. T.
Forthcoming in: Information Systems Research, 2021.

efl insights

The efl publishes the insights in the form of a periodic newsletter which appears two times a year. Besides a number of printed copies, the efl insights is distributed digitally via E-mail for reasons of saving natural resources. The main purpose of the efl insights is to provide latest efl research results to our audience. Therefore, the main part is the description of two research results on a managerial level – complemented by an editorial, an interview, and some short news.

For receiving our efl insights regularly via E-mail, please subscribe on our homepage www.eflab.de (> NEWS > SIGN UP EFL INSIGHTS) as we need your E-mail address for sending the efl insights to you. Alternatively, you can mail your business card with the note "efl insights" to the subsequent postal address or send us an E-mail.

Prof. Dr. Peter Gomber
Vice Chairman of the efl – the Data Science Institute
Goethe University Frankfurt
Theodor-W.-Adorno-Platz 4
D-60629 Frankfurt am Main

insights@eflab.de

Further information about the efl is available at www.eflab.de.

The efl – the Data Science Institute is a proud member of the House of Finance of Goethe University, Frankfurt.
For more information about the House of Finance, please visit www.hof.uni-frankfurt.de.

THE EFL – THE DATA SCIENCE INSTITUTE IS AN INDUSTRY-ACADEMIC RESEARCH PARTNERSHIP BETWEEN FRANKFURT AND DARMSTADT UNIVERSITIES AND PARTNERS DEUTSCHE BÖRSE GROUP, DZ BANK GROUP, FINANZ INFORMATIK, AMERICAN EXPRESS, DEUTSCHE LEASING, AND FACTSET LOCATED AT THE HOUSE OF FINANCE, GOETHE UNIVERSITY, FRANKFURT.

For further information please contact:

Prof. Dr. Peter Gomber
Vice Chairman of the
efl – the Data Science Institute
Goethe University Frankfurt
Theodor-W.-Adorno-Platz 4
D-60629 Frankfurt am Main

Phone +49 (0)69 / 798 - 346 82
E-mail gomber@wiwi.uni-frankfurt.de

Press contact
Phone +49 (0)69 / 798 - 346 82
E-mail presse@eflab.de

or visit our website
http://www.eflab.de