

Abstracts of the Swiss Track Presentations

Contents

- 1 Sentiment Analysis for a Swiss Gig Platform Company
- 2 Chatbots usability - Dos and don'ts
- 3 Neural Machine Translation for Increased Human Translation Efficiency at Migros Bank
- 4 QUAND - Quantitative Need & Demand Modeling from Reviews
- 5 Automated Data Extraction from Documents using Machine Learning
- 6 Summarization of German texts
- 7 Text-to-Speech (TTS) for Seven Swiss German Dialects
- 8 Success of ASR (automatic speech recognition) in the field of subtitling
- 9 Building an inductive coding service for employee feedback
- 10 Automatic News Generation & Reader's Comments Processing
- 11 Neural Text Normalization with Adapted Decoding and POS Features
- 12 How Patient Insights from Social Media improve a Conceptual Disease Model
- 13 DeepDTI: Identification of Drug-Target Interactions using Deep Learning Techniques
- 14 Quality Assessment of Automatic Speech Recognition Systems
- 15 Matching Human Competences with Semantic Search
- 16 Hyperparameter Tuning for Deep Learning in Natural Language Processing
- 17 PLACAT: A user-friendly question answering system for smart speaker devices
- 18 News from dialektkarten.ch
- 19 Email Anonymization in the insurance Industry
- 20 Semantic systems and visual tools for analyzing domain-specific text corpora and communication channels
- 21 The PsyMine Project: Outcomes and Insights
- 22 MedMon: social media analytics for an healthcare application
- 23 Multi-language multi-label email classification at SBB's customer support
- 24 Customer driven architecture for a Chatbot

1 Sentiment Analysis for a Swiss Gig Platform Company

Ela Pustulka-Hunt and Thomas Hanne

We work with a Swiss Gig Platform Company to identify innovative solutions which could strengthen its position as a market leader in Switzerland and Europe. The company mediates between employers and employees in short term work contracts via a platform system. We first looked at the business processes and saw that some process parts were not being controlled by the company, which is now being remedied. Second, we analyzed the job reviews which the employers and employees write, and implemented a prototype which can detect negative statements automatically, even if the review is positive overall.

We worked with a dataset of 963 job reviews from employers and employees, in German, French and English. The reviews have a star rating (1 to 4 stars), with some discrepancies between the star rating and the text. We scored the reviews manually as negative or other, as negative reviews are important for business improvement. We tested several machine learning methods and a hybrid method from Lexalytics. Training on multilingual Twitter data delivered an accuracy of 77%. The Lexalytics solution had an accuracy of 80%. Using manually annotated data, with cross validation, produced an accuracy of 86%, with logistic regression and SVM performing best. We are now designing a sentiment dashboard. Validating the findings with a dataset of 2.5K of manually scored reviews, produced an accuracy of 90%. Ongoing work is addressing review topics.

References: Pustulka-Hunt, Ela, et al. "Multilingual Sentiment Analysis for a Swiss Gig." 6th International Symposium on Computational and Business Intelligence (ISCBI 2018), Basel, Switzerland. 2018. Pustulka-Hunt, Ela, Rainer Telesko, and Thomas Hanne. "Gig Work Business Process Improvement." 6th International Symposium on Computational and Business Intelligence (ISCBI 2018), Basel, Switzerland. 2018. (to appear in IEEE Explore)

2 Chatbots usability - Dos and don'ts

Christoph Süess

More than twenty-five years ago the first web pages popped up in the internet. The process to deploy a web site was rather technical and topics like usability and user experience did not get a lot of attention. This changed a lot. Nowadays is all about user-centered design.

With chatbots the history repeats itself. The first chatbots were created by programmers who were fascinated by the technical possibilities. But how can we develop user-friendly chatbots?

In this presentation Christoph Süess will explore popular Swiss chatbots related to different industries such as banking, insurance, tourism and mobility. Join the talk, take a look at chatbot conversations and experience smart and less clever practices. Learn how wording, dialogue structure, fallback strategies, the use of quick answers and other technics can impact the user experience.

This talk will arm you with a bag of tips and tricks to build chatbots which boost the user engagement, which appear intelligent and trustworthy and that help you reach your business goals.

3 Neural Machine Translation for Increased Human Translation Efficiency at Migros Bank

Stephan Wick and Samuel Läubli

Migros Bank translates more than one million words per year from German to French and Italian. To support their internal team of human translators, the bank decided to develop an in-house neural machine translation system as the use of online services was ruled out for quality and security reasons. This system has been operational on Migros Bank's own infrastructure since mid-2018, meaning that any new email, website or marketing text can be translated automatically without ever leaving the company's network.

In this presentation, we summarise lessons learnt from both a business and technical perspective. We also present results from a blind comparison of the Migros Bank system with DeepL, and results from a productivity study with the internal translators. The latter shows that human translators are 35% faster when using Migros Bank's machine translation system.

4 QUAND - Quantitative Need & Demand Modeling from Reviews

Ritavan -, Sameh Metias and Harald Koppen

We present a solution based on Natural Language Understanding algorithms that analyses online reviews and discussions to quantify opinion and provide actionable insights for improving product or service features to better understand customer opinion. The text is usually short, noisy and unstructured thus most standard methods don't function in this setting. Additionally, for the solution to be adaptable and scalable, all methods involved have to be essentially non-supervised or at worst semi-supervised. We thus demo completely unsupervised topic modelling, as well as weak aspect-based sentiment and opinion quantification. Our solution is specifically targeted at all forms of retail settings. Relevant industries are thus e-commerce, consumer electronics, travel, automotive, banking, insurance etc.

Our current solution is focused primarily on the DACH region as it is available in German and English. Most of the datasets we currently use are also based on companies that have a strong DACH focus.

5 Automated Data Extraction from Documents using Machine Learning

Aaron Richiger and Sara Wick

About 80 percent of all business-relevant data is unstructured. Given the variety in layouts and changing contexts of information in a document, data fields or texts are still manually copied for further processing. This is time-consuming, error-prone and costly. Bringing semi-structured and unstructured data into a machine-readable format is therefore an ongoing challenge and prerequisite also in the field of text analytics.

Over the last years, we developed a document extraction platform – MINT.extract – that allows us to access relevant data (text, images, etc.) from any type of document in a highly efficient and flexible way. In this talk, we present a customer case to process purchase orders using machine learning. Besides automated extraction of information like order number, article number, ordering date, and the number of items ordered, we also perform validation steps. The system runs as a service 24/7 and is accessible from all our client's global locations. As a further benefit, the learning system is highly scalable given the small set of training data required to get the accuracy scores above 95%. We will share some technical insights on how we achieved this. Finally, an intuitive user interface allows users to train their own learning system independently and the architecture enables a seamless integration with existing ERP systems. Naturally, this approach can be applied to insurance policies, invoices, official documents or entire book series as well as other document types.

6 Summarization of German texts

Dominik Frefel

Most research on automatic summarization is limited to the English language. My master thesis at FHNW explores the possibility to apply this knowledge to German and proposes a deep learning algorithm for summarizing German texts. The developed summarization algorithm has two core components: sentence extraction and abstractive summarization. The sentence extraction module predicts the usefulness for each sentence of a given text. The second component takes the identified most useful sentences as input to create an abstractive summary. The architecture is a pointer generator network [1] combined with the Transformer model [2]. Both components are trained simultaneously.

The model is trained with texts extracted from the German Wikipedia. To create a German summarization corpus, articles are identified which have an introduction that summarizes the rest of the text sufficiently well. These introductions are used as reference summaries.

Below, an example is given of a generated summary and an excerpt of the corresponding source text. Source: Beaconsfield liegt im Westen der Île de Montréal in der Region West Island, am Nordufer des Lac Saint-Louis. Die Gemeinde grenzt im Westen an Baie- D'Urfé, im Nordwesten an Sainte-Anne-de-Bellevue, im Norden an Kirkland und im Osten an Pointe-Claire. Das Stadtzentrum von Montreal ist rund 22 Kilometer entfernt. Die erste Siedlung auf dem Gemeindegebiet entstand 1698. . . . Mehrere Buslinien der Société de transport de Montréal stellen Verbindungen mit den Nachbargemeinden her.

Generated summary: Beaconsfield ist eine Stadt im Südwesten der kanadischen Provinz Québec. Sie liegt in der Region West Island, etwa 22 Kilometer nördlich von Montreal.

[1]: <https://arxiv.org/pdf/1704.04368.pdf> [2]: <https://arxiv.org/pdf/1706.03762.pdf>

7 Text-to-Speech (TTS) for Seven Swiss German Dialects

Christof Traber, Schamai Safra, Bleicke Holm, Dominic Schnyder and Philipp Lichtenberg

SlowSoft GmbH develops text-to-speech (TTS) components for Swiss German dialects. By mid-2019, TTS for seven major dialects (GR, ZH, BE, LU, SG, BS, VS) will be available, each at first with a single voice (three male, four female voices).

Typical applications of Swiss German TTS include speaking aids (e.g., for ALS patients), language learning tools, chatbots, robots, and digital assistants. Even for applications that finally use recorded speech as output (e.g., IVR systems), TTS may provide an excellent means to create prototypes during the development phase of the application.

Whereas the voice generation component (the "back end") of Swiss German TTS is comparable to that of other languages, the NLP component (the "front end") is extremely challenging due to the fact that there is no standardized way of writing Swiss German. A desirable choice of textual input for most TTS applications would therefore be to use Standard German as input, which, however, would require a full-fledged translation from Standard to Swiss German. A somewhat easier solution is the use of normalized Swiss German, i.e., a word-by-word representation of the Swiss German content by means of Standard German words. The currently adopted solution lies somewhere between these extremes.

The presentation describes the state of development of the Swiss German TTS system, with a focus on the challenges associated with the input text processing.

8 Success of ASR (automatic speech recognition) in the field of subtitling

Gion Linder and Jacqueline Blaser

Subtitling is a labour-intensive activity and thus expensive. Therefore, there is a strong pressure to replace this activity with automatic speech recognition and post-editing. But does it really work? Do the automatic subtitles meet the quality requirements? And how should the correction loop be designed so that the texts created with automatic speech recognition can be corrected efficiently? The subsequent proofreading process is almost as important as the quality of automatic speech recognition, since manpower for proofreading is the biggest cost factor.

9 Building an inductive coding service for employee feedback

Andrew Marritt

Large firms are increasingly using feedback as a way of collecting information on ideas on a wide-range of business and employee topics. Workometry, a 2018 Gartner Cool Vendor, is a service for large employees that enables them to quickly make sense of large volumes of multi-language employee feedback.

Feedback, often provided as sentence-fragments, needs to be understood in the context of the question. Decision makers have a strong preference for identified themes to themselves be reasonable answers to the question rather than noun-based topics. We therefore build inductive, question and organization-specific models aiming do this within a few hours. Our approach uses a human-in-the-loop to improve accuracy to perform a service with similar performance to human coders but at a fraction of the resource requirements.

In this presentation I discuss how we built and continually refine the service. I will discuss some of the challenges we overcame and how small changes in the survey employees are presented can improve (or decrease) the ease of building successful models.

10 Automatic News Generation & Reader's Comments Processing

Didier Orel, Titus Plattner and Marcel Blattner

In this talk, we will survey how Tamedia plans to benefit from recent technologies for Natural Language Generation to implement new and innovative methods for news generation. Tamedia is the largest private media group in Switzerland, owning more than 30 newspapers in the country : Tages Anzeiger, 20 Minuten, Sontagszeitung, Tribune de Genève, Berner Zeitung, Basler Zeitung, 24 heures, ...

Tobi is Tamedia's automated text project for vote results at a local level. On November 25th 2018, in just five minutes, we generated about 40.000 texts – with customized variations for all 2222 municipalities in two languages. Three months later we did the same, including even cantonal and local results.

These successful pilots for vote results demonstrated how we can address a long tail audience with localized news. These texts were generated with a template-based approach. Other media organizations, with Associated Press at the vanguard since 2014, already use this approach. But we think that a personalized distribution of these texts – with for example personalized push alerts or a tailored feed – will allow leveraging a large quantity of micro-audiences. The key success factors of these pilots were: the experienced journalists who crafted the templates; smooth user experience thanks to an elegant and clear front-end; and the measure of user feedback. Based on the positive outcome of the pilots, Tamedia plans to automate texts for sports results.

Tamedia is also enabling advanced models of text analysis to process the huge amount of comments readers write on the different digital channels. Goals are on one hand to detect toxic comment in order to support the moderation process speed up the review time, and on another hand to surface the most interesting comments and extract a snippet from them (see <https://twitter.com/tagicommments>).

11 Neural Text Normalization with Adapted Decoding and POS Features

Tatyana Ruzsics, Massimo Luseti, Anne Göhring, Tanja Samardzic and Elisabeth Stark

The task of text normalization aims at bringing non-canonical language, coming from speech and social media, to a standardized writing. This task is especially important for languages such as Swiss German, with strong regional variation and no written standard. In this work we propose a novel solution for normalizing Swiss German WhatsApp messages using the encoder-decoder neural machine translation (NMT) framework. We enhance the performance of a plain character-level NMT model with the integration of a word-level language model and linguistic features (POS tags). The two components are intended to improve the performance by addressing two specific issues. The former targets the fluency of the sequences predicted by NMT: it corrects a sequence which is not a proper word, despite being a likely sequence of characters. In addition, this modification targets the frequent cases where a contracted form corresponds to multiple normalized words, e.g. the word *kömmmer* 'we can' is mapped to the normalization form *können wir*. The latter component, the addition of POS tags, aims at resolving cases of word-level ambiguity. For example, the ambiguous input word *Lüüt* can be normalized as the noun *Leute* 'people' or *läuten* 'to ring' when used as a verb. Our systematic comparison shows that the proposed model improves over the best previous solution. A thorough analysis of the compared systems' output shows that our two components produce indeed the intended, complementary improvements.

Intended audience: Our intended audience includes both researchers and professionals interested in applying our resources in their own work.

12 How Patient Insights from Social Media improve a Conceptual Disease Model

Mathias Leddin, Raul Rodriguez-Esteban and Juergen Gottowik

The popularization of disease forums and microblogs has enabled the development of social media mining for biomedical research. Both regulatory agencies and pharmaceutical companies recognize that social media content provides valuable and unfiltered insights into the perspectives of patients and caregivers. Such insights can be key in the development of patient-centric outcome research, which seeks to strengthen the role of patient opinions in the selection of clinical trial outcomes instruments. However, because social media content is Real World Data, it poses unique challenges in comparison to scien-

tific literature mining. Some of these challenges are: higher prevalence of non-English content, use of colloquial and lay language, abundance of noise and junk content, and source format variability.

This presentation will exemplify how social media mining can inform clinical trial design by a thorough analysis of the patient perspective in Parkinson's disease. An analysis of early symptoms and impacts on quality of life described by these patients online was shown to improve a conceptual disease model based on the scientific literature and clinical experts. The results were confirmed by patient focus groups. The conceptual disease model was then used to define patient relevant clinical endpoints in current Roche trials. Thus, our work proved that social media mining is a valid data source in patient-centric outcome research.

13 DeepDTI: Identification of Drug-Target Interactions using Deep Learning Techniques

Hatem Ghorbel, Fabrizio Albertetti and Jérôme Moret

Drug target interaction (DTI) is crucial in drug research as it facilitates the process of drug discovery, drug side-effect prediction and drug repositioning (Ding et al. 2014). Most of in silico approaches are particularly oriented by the genome and chemical structure similarities, more than by pharmacological information. Their databases often fail in coverage and completeness (e.g. only known and approved drugs are included) (Dai & Zhao, 2015).

We suggest a text mining approach as a complement to the existing approaches to predict unknown DTIs that cannot not be deduced by structural information described in the conventional resources. We focus mainly on extracting interactions from MEDLINE abstracts.

Our work defines a deep learning convolutional architecture to predict DTIs. It uses natural language features from text such as drug/target entities, part-of-speech and dependency graph. We compare our approach to the BioCreAtIvE challenge recently organized to promote the same task (2017-2018). Their corpus provides drug/target entities and interactions labeled as five class representing groups of interactions (e.g. activator, inhibitor, etc.). Our model achieves a micro averaged f1-score of 0.63 which equals the best challenger. Our next goal is to use transfer learning approaches to automatically build features and obtain richer semantics. Moreover, a comparison with attention based recurrent networks is in progress.

14 Quality Assessment of Automatic Speech Recognition Systems

Fabian Germann, Malgorzata Anna Ulasik and Mark Cieliebak

Tech companies developing Automatic Speech Recognition (ASR) systems are in constant pursuit of achieving the lowest possible Word Error Rate. Google with their 4.9% WER is currently in the lead, followed by Microsoft's 5.1% and IBM's 5.5%. The range of ASR systems is wide: from open-source to closed-source tools, offering APIs or requiring more complex integration, achieving better or worse performance. Various aspects can be considered when selecting a proper solution but certainly accuracy is the key factor. Our study provides an in-depth accuracy comparison of the current state-of-the-art solutions, including Google Speech-To-Text, IBM Watson, Microsoft Azure, Mozilla DeepSpeech, Amazon Transcribe and CMUSphinx. The analysis is based on transcriptions of several well-known speech corpora such as Timit, Switchboard, LibriSpeech or CommonVoice. This results in over 50 ASR System - Speech Corpus pairs constituting a solid foundation for the comparison. The accuracy is evaluated based on the Word Error Rate. The results of this study will enable practitioners to select the "best" solution based on their needs.

15 Matching Human Competences with Semantic Search

Kurt Wehrli, Massimo Lusetti and Tanja Samardzic

In our fast-evolving society, home services potentially represent a valuable resource to foster friendships, enrich lives, acquire know-how, share competences and discover capabilities. The goal of the Symsala project is to develop a web platform to serve as online market place of individual skills and talents. This exchange requires matching customers' needs with skills and competences offered by providers. The

Symsala association and the University of Zurich are collaborating on the core NLP task of the platform, which consists in matching a customer's query with the corresponding skill in the competence data base. Similarly, a description of the skills and services that a provider can offer is matched with the relevant skills in the data base, thus creating a provider's profile. The matching task is carried out using word embeddings, where a word is viewed as a vector in a multi-dimensional space. Word embeddings allow us to compute distance and similarity between words and phrases by means of specific metrics, such as cosine similarity. We run experiments with different pre-trained word embeddings in both English and German, and explore various methods that can be applied to compute the similarity between word sequences, for example word mover's distance and sentence embeddings obtained from the single word embeddings. The outcome of these experiments provides a basis for a future innovative semantic search solution in the context of the Symsala platform.

Intended audience: professionals and general public.

16 Hyperparameter Tuning for Deep Learning in Natural Language Processing

Ahmad Aghaebrahimian

Although Deep Learning has advanced a lot through past several years, it still seems like a black art for many people mostly due to the fact that obtaining consistent and good results from a deep architecture requires optimizing many parameters. Hyperparameter tuning is an essential task in deep learning which can lead to significant changes in the network performance. This talk is the essence of approximately 3000 GPU hours on optimizing a network for a text classification task on a wide array of hyperparameters. Word embedding types, word embedding sizes, word embedding updating, character embedding, character embedding sizes, deep architectures (CNN, LSTM, GRU), optimizers, gradient control, classifiers (Softmax, Sigmoid, CRF), dropout, deep vs. wide networks, pooling, and batch sizes are the hyperparameters studied in this work using a grid search scheme. I will talk about the most critical parameters and the insight behind them that researchers can modify or prioritize in a deep architecture to get the best performance with the least effort on the part of humans and the machine.

17 PLACAT: A user-friendly question answering system for smart speaker devices

Gabriel Luthier and Andrei Popescu-Belis

The goal of the PLACAT project is to combine the conversational capacities of chatbots, which can be trained on large amounts of non-directed dialogue, with the question answering (QA) capability of dedicated systems. The result is demonstrated as a virtual assistant on a commercial smart speaker, using its automatic speech recognition capabilities. The virtual assistant includes three key technical components, two from the state of the art and one innovative.

The key innovation is a dialogue controller that directs input utterances, depending on their recognized dialogue act, either to a robust chatbot (if the utterance is labeled as interaction-related, such as greetings or chit-chat) or to a powerful question answering system (if the utterance is recognized as a request for information). The controller is a classifier trained on a database composed of questions from chat corpora and from QA ones. The chatbot uses recent models based on reinforcement learning trained on chat corpora, while the QA system uses the recent BERT representations of words in context (from a Transformer neural network) trained on a large QA database (SQuAD).

The demonstrator uses the Google Home speaker. It forwards spoken actions to the controller on the DialogFlow platform, which then connects to our servers for each of the dialogue components. The PLACAT platform is intended as a showcase for the development of commercial virtual assistants endowed with a strong capacity for question answering, particularly over corporate documents such as technical manuals, but at the same time able to sustain user-friendly conversations.

18 News from *dialektkarten.ch*

Yves Scherrer

The web site *dialektkarten.ch* has been online since 2014. It presents various interactive visualisations of the Swiss German dialect atlas SDS (Sprachatlas der deutschen Schweiz). Besides a set of digitized feature maps, it also provides results of dialectometric analyses as well as prototypes of machine translation and dialect identification systems for the entire German-speaking area of Switzerland. In this paper, we present three recent developments of this web site: - The mapping backend has been changed from Google Maps to Leaflet. - The data set has been extended with additional digitized SDS maps. - The same interactive backend has been applied to new sister projects which present interactive visualisations of French, Italian and Romansh dialect data. These sister projects have been realized in collaboration with the Salzburg dialectometry group under Prof. Hans Goebel and are based on the ALF (Atlas linguistique de la France) and AIS (Atlante linguistico ed etnografico d'Italia e della Svizzera meridionale), respectively. Taken together, the SDS, ALF and AIS dialect maps cover all four linguistic areas of Switzerland. We will also provide an outlook on future projects.

19 Email Anonymization in the insurance Industry

Anastasios Zouzas, Theus Hossmann, Benjamin Theunissen and Christian Müller

Electronic mail is arguable the most frequently used method of internal communication within various enterprises as well as the insurance sector. Several business processes within the insurance industry are orchestrated using emails which promotes email analytics as an important task for data scientists / business analysts. In general, access to such highly sensitive data across several departments within an organisation is not permitted without the removal of sensitive / personal information (anonymization). In this talk, we will discuss a randomised replacement approach on email anonymization that is designed and developed on real data within the Swiss insurance industry. In more detail, we will demonstrate how to engineer such a solution based on open-source implementations of named entity recognition that is enhanced with external and internal lexicons.

20 Semantic systems and visual tools for analyzing domain-specific text corpora and communication channels

Albert Weichselbraun and Philipp Kuntschik

Web sites, News and social media represent important interfaces to consumers and stakeholders. The same is true for an organization's internal data such as documents, wikis and content stored in collaboration services.

Systematically identifying stakeholders and their communication channels has, therefore, the potential to yield insights into their views, provides means to identify unfolding trends at an early stage, and allows researchers and practitioners to gain a more comprehensive understanding of the analyzed domain and its dynamics.

The proposed presentation demonstrates – based on the Swiss Media Criticism project – how domain knowledge provided by social and communication scientists is combined with publicly available linked open data sources, natural language processing components and machine learning methods, to provide efficient tools for analyzing and understanding comprehensive, heterogeneous document collections. A data acquisition, text mining and knowledge enrichment pipeline acquires and selects relevant content that is then structured, annotated and contextualized based on the provided domain model. The enriched documents are stored in a contextualized information space.

A powerful visual web dashboard allows real-time analyses and visualizations of the created information space, yields information on unfolding trends and topics, and visualizes key entities such as stakeholders, their relations to each other and concepts that are associated with them.

21 The PsyMine Project: Outcomes and Insights

Tilia Ellendorff, Simon Foster and Fabio Rinaldi

We present outcomes of the PsyMine project, a 2-year collaborative project with the aim of identifying etiological factors of mental disorders in the text of research articles.

Mental disorders are one of the main contributors to the contemporary burden of disease and one of the major groups of disorders that cause disability, not only in Switzerland but also in the rest of the world. One priority of research in mental health is to discover etiological factors (underlying factors) of mental disorders. However, ongoing research in mental health faces the problem of having important information scattered across numerous textual resources. One of the main reasons for this situation is etiological factors can be covered by different areas, such as genetics, sociology, chemistry, and many more. Most relevant information concerning etiological factors is encoded in the unstructured written text of research articles of the various sub-domains. This fact makes it hard for researchers to gain an overview of previously discovered findings and therefore obtain a complete picture of a mental disorder and involved factors.

The PsyMine project, which is now completed, focused on several aspects of text mining for psychiatric disorders: availability of task-related data, data annotation, terminology and, most importantly relation/event-extraction to identify etiological events in literature. The applied methods present ways towards automatizing the overall process of collecting information from unstructured, written text in the domain of mental health and converting it into an easily accessible, structured database. We present the overall project as well as main insights from the different involved aspects.

22 MedMon: social media analytics for an healthcare application

Fabio Rinaldi, Philipp Kuntschik, Gottowik Jürgen, Mathias Leddin, Raul Rodriguez Esteban, Albert Weichselbraun, Tilia Ellendorff, Nico Colic and Lenz Furrer

We present MedMon, a collaborative project supported by InnoSuisse, focused on the analysis of social media to capture the exchange of biomedical information among providers, patients, scientists, pharmaceutical companies and other healthcare stakeholders.

The project aims at developing a text analytics platform specifically tailored to social media and to be used within a major pharmaceutical company. Two application scenarios under consideration are described below:

1. Clinical trials only capture parameters predefined in protocols. Through social media it is possible to gather patient discussions on disease burden, co-morbidities, quality of life and activities of daily living to identify aspects not covered by trial protocols to reduce potential study burden for patients.

2. Emerging information relevant to different disease areas needs to be monitored, filtered and aggregated, however patient blogs, social media and news channels cannot be queried systematically with existing tools. MedMon aims at providing more efficient means to query these data sources. Using social media analysis to gain improved patient insights has the potential to save significant amounts of resources in the process of drug development, and improve patient outcomes measures.

The project partners have developed a fully functional prototype of the analytical platform and are researching technical advancements to refine the accuracy of the delivered results.

23 Multi-language multi-label email classification at SBB's customer support

Nadja Keidel and Marlyse Lichsteiner

Since more than 15 years the inboxes of the agents at the customer service center are filled with emails asking for advice, support or providing feedback for the Swiss Federal Railways' services and products. In addition to answering the client's email, it is also crucial for the team to ensure that problems and suggestions for improvement are forwarded accordingly to the responsible person within SBB to make change and improvement hap-pen. However, the volume of customer emails increases year by year. Thus, the customer service center depends on an efficient solution to support the quantitative and qualitative analysis of the incoming emails. As one of the first operative machine learning projects at SBB, we implement a customized multi-language multi-label email classification to break down the incoming

traffic into its main topics. Moreover, we empower the users to leverage their individual quantitative and qualitative analyses of the customer feedbacks by providing insightful visualizations and a base for efficient self-service reporting. In close collaboration with our test user group, we develop and adapt the dashboard according to their needs. We would like to share our general setup of components, architecture and implementation approach combined with anecdotes of surprises and challenges that came up from both business as well as technical perspective. For instance, with respect to the creation and enrichment of the ground truth, dealing with the multi-language scenario and quite some imbalances in the dataset.

24 Customer driven architecture for a Chatbot

Laura Gander, Antonella Bolt and Ursula Stäuble

1.26 million of passengers are travelling in Switzerland with the SBB every day. The customer service before, during and after the journey is therefore more than important. Our customers need a quick and handy way to order tickets, reserve seats and obtain information.

The volume in our customer service center increases year by year. Moreover, a big share of the requests are similar to each other. Accordingly, to keep the volume on a level, reduce cost and let our employees focus on more complex request, we thought about a chatbot which is a highly qualified and advanced product assistant to help our customers 24/7.

We started a first pilot phase with a basic question-answer bot. The goal was to explore whether our customers would use a chatbot at all and to learn about the kind of questions they would ask. In order to find gather more information about our customer we conducted qualitative interviews to get insights in what customers know about using chatbots and what expectations they have.

With the gathered insights of the question-answer bot we created a totally new structured different bot. We revised both architecture and design of the bot and included machine learning modules for more “intelligence”. To a certain extent, the bot is now able to capture context and sentiments and recognize when it is time to hand over to an agent in a live chat.

The iterative approach allowed us to integrate customer continuous feedback and leveraging the bot to react more appropriate to the customer’s requests. While still being in the process of further improvements, we would like to take this opportunity to present you our findings also in comparison with an almost completely dialog guided chat bot (rule-based). What did work so far, and what did not?

- how to gather insights in user behaviour/interaction with a chatbot - which use cases are suitable for a chatbot - why it is so important to give the bot emotions and how to detect them