Machine Learning models for the automation of accounting process

Chiara Bardelli

CompMat22
Spring Workshop 2022
Università di Pavia, 17/03/2022

Outline

- Introduction
- Phase 0: the research project
- Phase 1: testing user phase
- Conclusions and Q&A

Introduction

DATEV

About me

NOW

Data scientist in Datev S.p.a.

2021

PhD in Computational Mathematics and Decision Sciences at University of Pavia

2016

Master degree in Mathematical Engineering at Politecnico di Milano

About Datev

Datev is a software company which provides softwares and IT services to facilitate officers in the accounting process.

Datev sponsored the scholarship of my PhD in Computational Mathematics at University of Pavia.

Why a software company is interested in a collaboration with the University of Pavia and decided to sponsor a scholarship?

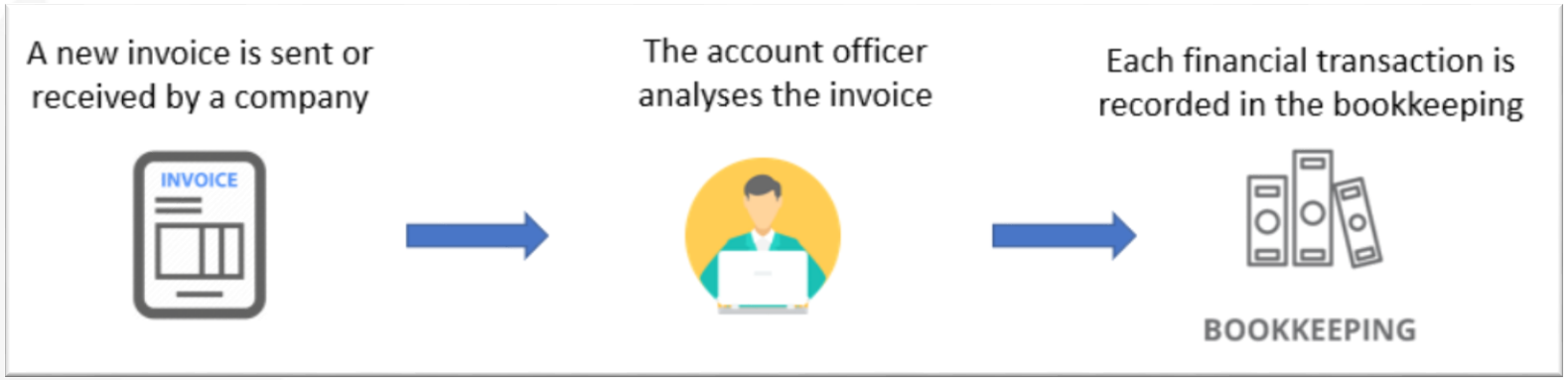
DATEV

Let’s talk about accounting process...

The accounting process includes different steps to record the financial transactions of a company. The transactions are collected, identified, ***classified****, summarized and recorded in the bookkeeping journal so that the financial position of the business can be known after regular intervals of time.*

Let’s talk about accounting process...

Part of the accounting process is a repetitive task which could be a good candidate for the introduction of **an intelligent classification system** based on historical data.



Machine Learning models can be used for in the creation of bookkeeping journal?

About this project

Machine Learning models can be used for in the creation of bookkeeping journal?

Datev had no previous experience with Data Science projects and Machine Learning models.

Datev was not sure about the success of the project.

The complexity of the problem and data needed a detailed research phase.

To answer this important question Datev decided to start a collaboration with University of Pavia, providing a scholarship for a PhD in Computational Mathematics and Decision Sciences.

DATEV

About this project

Phase 0:
Research phase

- Is it possibile to apply ML models to accounting data for our purpose? Results are good enough?

People involved in the project:

- Me
- Francesco and Mattia (.NET developer)
- Alessandro (tech leader)

Phse 1:
user testing phase

- Is it possibile to engineer the whole process? Results satisfy the requirements of software users?

Phase 2:
integrate the new feature in the software

- Design a suitable infrastratcure to provide a new service based on ML models to software users

DATEV

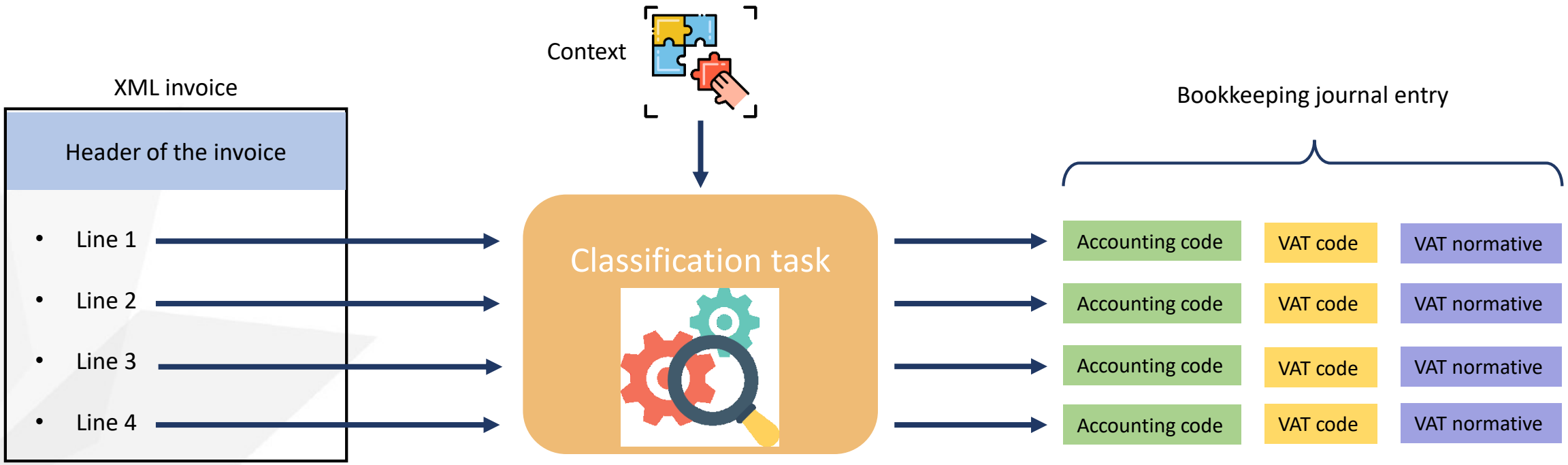
Phase 0: the research project

DATEV

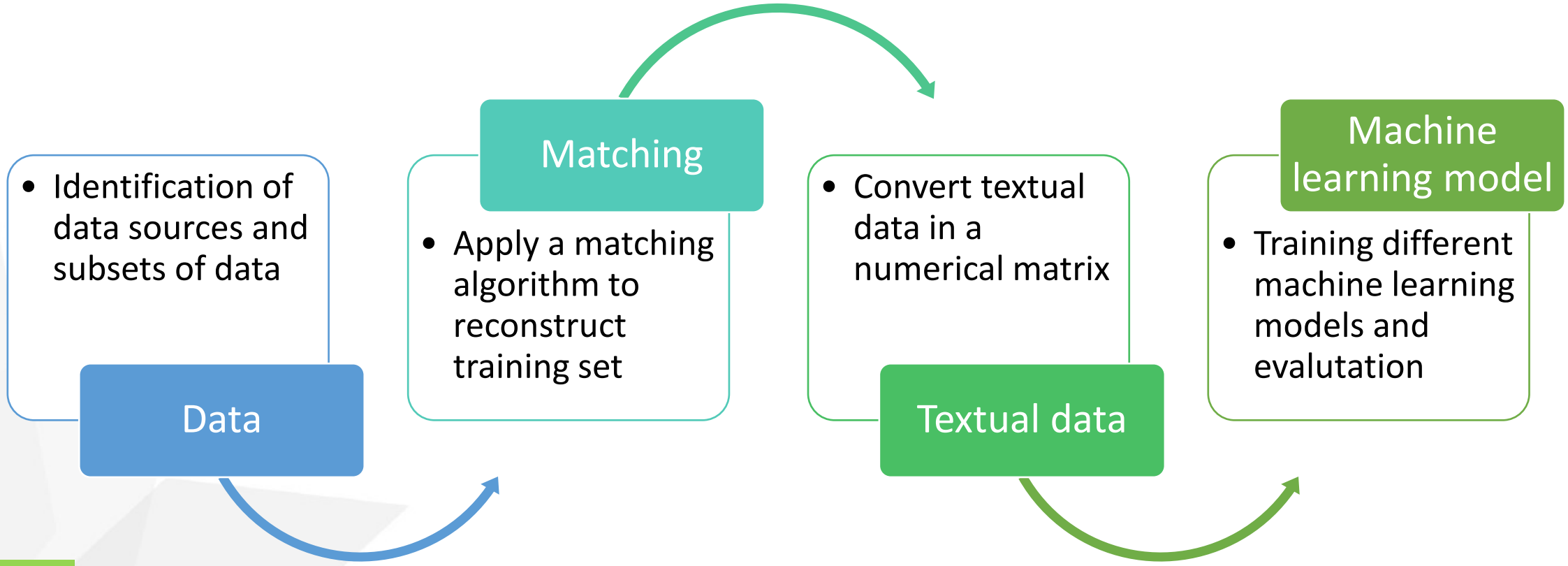
Domain of the problem

The bookkeeping journal entry is based on:

- Information contained in the electronic invoice
- Context in which the invoice has been processed (i.e. who are the sender and the receiver and which are their main occupations)



Research project



Data sources

The bookkeeping journal entry is based on:

- Information contained in the electronic invoice
- Context in which the invoice has been processed (i.e. who are the sender and the receiver and which are their main occupations)

Since January 2019, in Italy, the electronic invoice is mandatory. Each invoice is structured in a specific XML file with a fixed schema and easy to access.

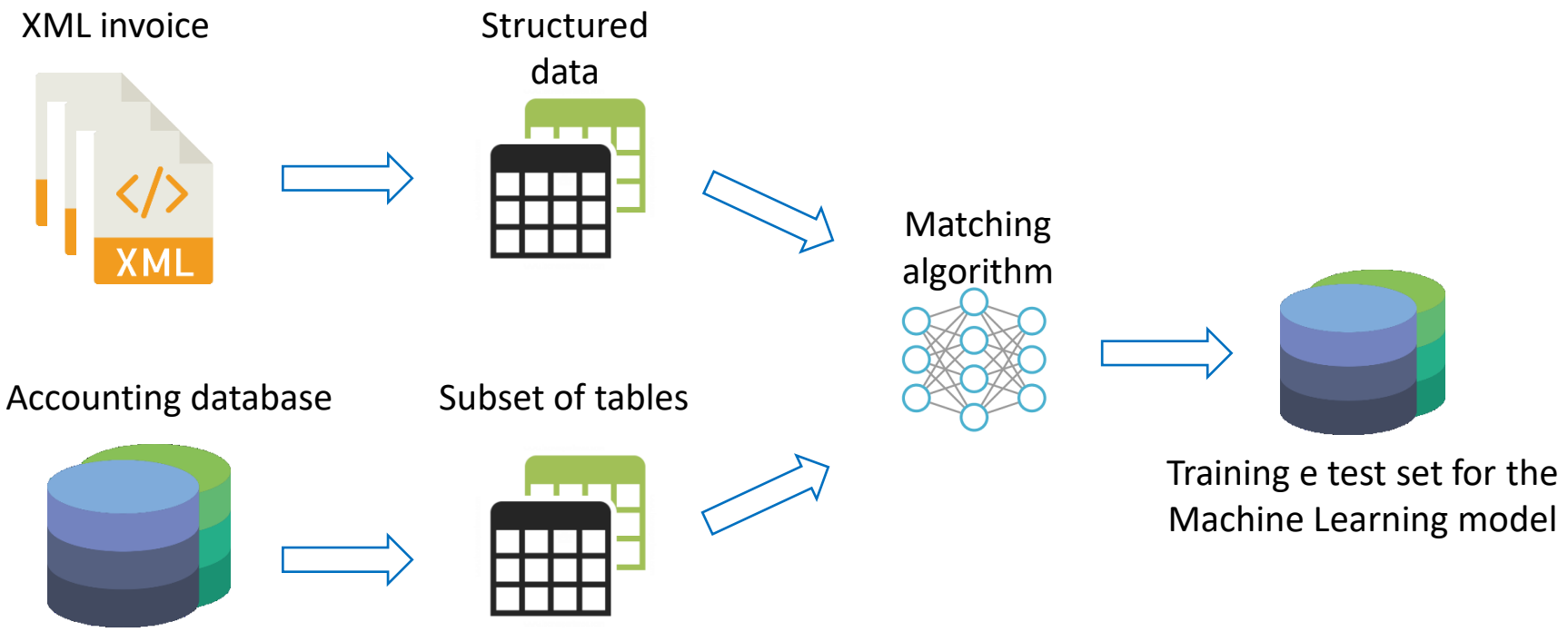
The accounting database includes all the history of the accounting process for a specific customer. Each invoice is represented as a single bookkeeping journal entry.

Data has been selected thanks to the collaboration with experts of accounting in Datev

Data extraction process

Information contained in the invoice

Context of the invoice



A matching algorithm is necessary to reconstruct the association between lines of invoice and bookkeeping journal entries.

Machine Learning models

Preprocessing step:

- Elaborate and create new groups of variables
- Clean textual data
- Apply a model to convert text in numical matrix
  - Bag Of Words
  - Word2Vec

Tune different classification algorithms and evalute them using cross validation

- Random Forests
- AdaBoost
- Multi layer perceptron
- …

Metrics to evaluate:

- Macro and average Recall, precision and F1-score

DATEV

Results

Table 2. Mean cross validation precision and f1-score of the classifiers which predict account codes (received invoices). Standard deviation is reported in parentheses.

	Dataset 1		Dataset 2	
	Precision	F1-score	Precision	F1-score
RF + BoW	83.2% (±1.5%)	82.6% (±1.4%)	86.7% (±0.6%)	86.4% (±0.8%)
RF + Word2Vec	87.4% (±1.7%)	86.9% (±1.5%)	90.0% (±0.5%)	89.8% (±0.6%)
AdaBoost + BoW	68.6% (±5.0%)	67.7% (±4.7%)	76.2% (±4.5%)	77.0% (±3.6%)
AdaBoost + Word2Vec	73.6% (±4.3%)	72.8% (±4.3%)	79.6% (±3%)	79.5.2% (±2.8%)
MlP + BoW	80.8% (±1.4%)	81.2% (±1.5%)	83.2% (±1.6%)	83.9.1% (±1.4%)
MlP + W2V	79.5% (±1.2%)	80.0% (±1.0%)	83.1% (±1.0%)	84.0% (±1.1%)

Table 4. Mean cross validation precision and f1-score of the classifiers which predict VAT codes (received invoices). Standard deviation is reported in parentheses.

	Dataset 1		Dataset 2	
	Precision	F1-score	Precision	F1-score
RF + BoW	93.4% (±0.8%)	93.7% (±1.0%)	92.2% (±0.9%)	91.8% (±1.1%)
RF + Word2Vec	94.8% (±0.5%)	94.6% (±0.5%)	94.3% (±0.6%)	94.2% (±0.6%)
AdaBoost + Bow	92.6% (±0.6%)	92.8% (±0.8%)	89.9% (±1.0%)	90.5% (±0.8%)
AdaBoost + Word2Vec	94.9% (±0.5%)	94.9% (±0.6%)	94.2% (±0.5%)	94.1% (±0.5%)
MlP + BoW	92.8% (±1.0%)	97.2% (±0.8%)	91.4% (±0.6%)	91.8% (±0.5%)
MlP + W2V	95.4% (±0.3%)	95.0% (±0.4%)	94.3% (±0.5%)	93.8% (±0.6%)

Different Machine Learning models have been compared on 2 real datasets provided by 2 different tax consultants.

Precision and other indexes have been computed to evaluate the predictive performances of different models.

These promising results have encouraged the following phase of the project.

DATEV

Phase 1: testing user phase

DATEV

Aims of the testing user phase

Why design an experimental phase?

- Test the architecture and the flow of the procedure

Architecture

- Extend the sample data thanks to the participation of different tax consultants

Data

- Collect statistics and results for each tax consultant to verify the stability of the process and the accuracy of the proposals

Accuracy

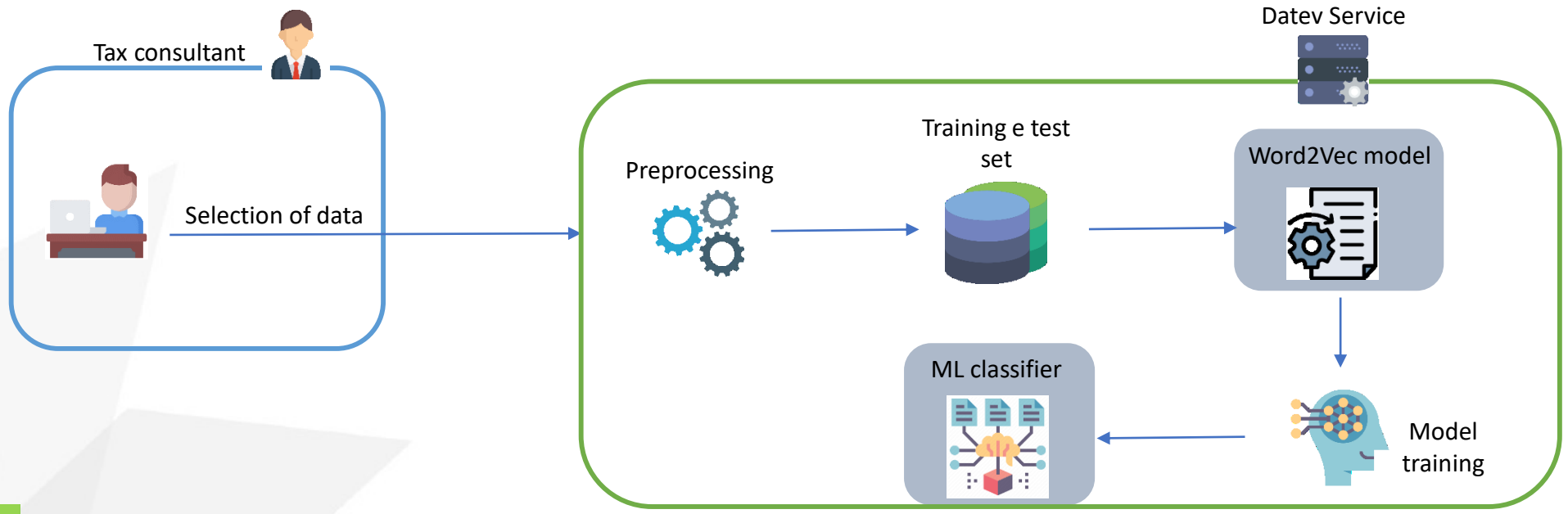
- Understand if data selected for the ML models are sufficient to identify the accounting rules and propose accurate bookkeeping entries

Input variables

DATEV

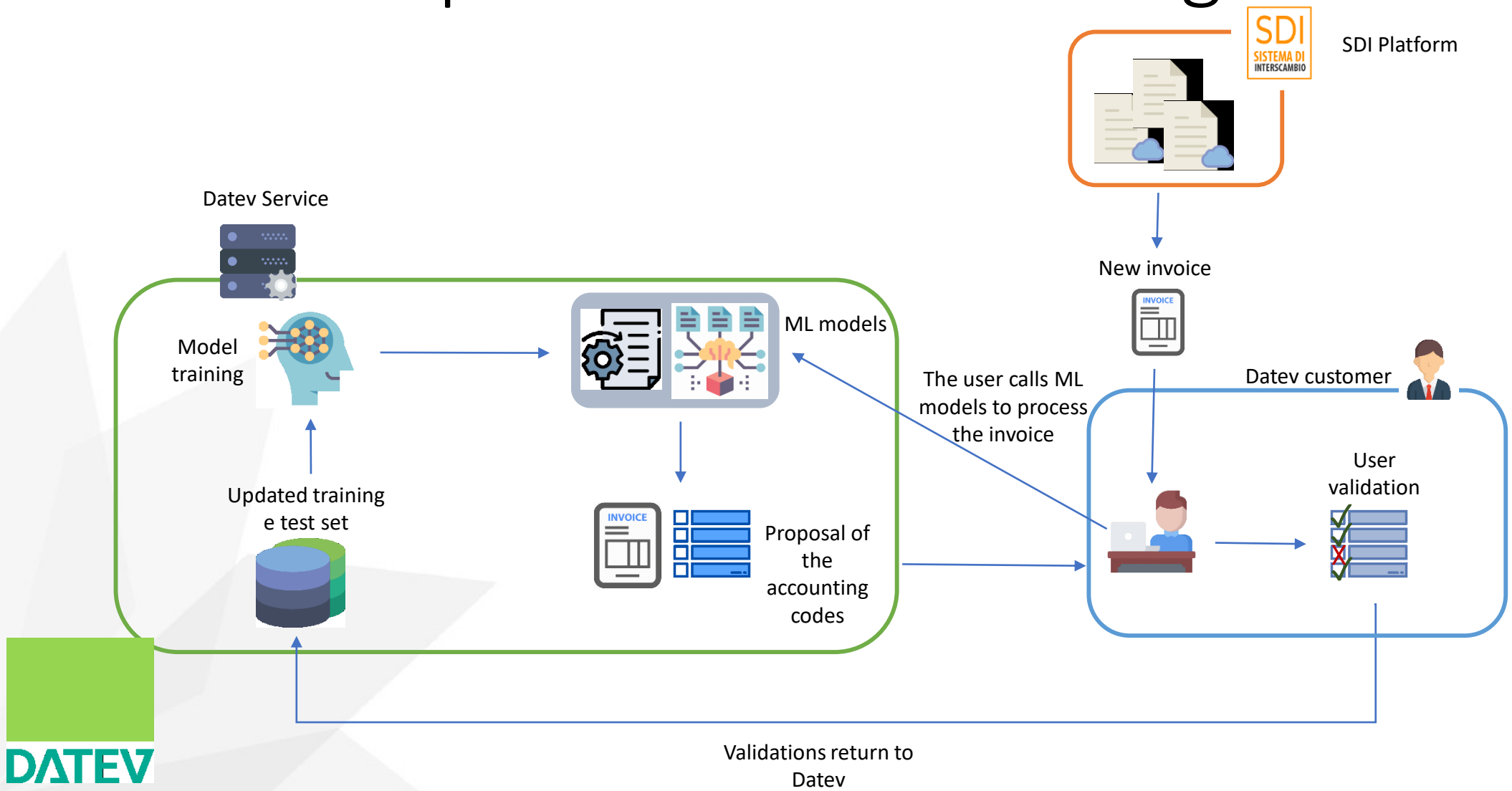
Integration into accounting process - Training phase

A Machine Learning classifier is trained on historical data to learn accounting rules.



Word2Vec model and ML algorithm are saved and used in the next phase.

Integration into accounting process - Prediction phase and re-learning



Experimental phase

How to design a testing user phase?

A new experimental feature (Automatic Accounting) was included in Datev software. The feature was enabled only for a restricted number of tax consultants previously contacted by phone.

The new experimental feature consists of two different components:

A wizard for the selection of clients which a tax consultant would like to test.

A new window containing the results and the proposals of the ML algorithm after importing an electronic invoice.

Challenges and problems

- ***Issues related to the contact to accounting firms***

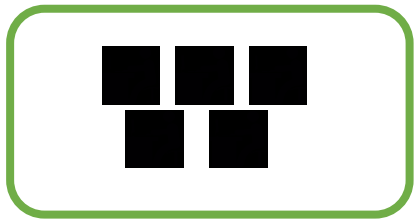
  The tax consultants selected for the experimental were required to send their feedbacks about the proposals of ML algorithms. Unfortunately, they were not so active in providing their feedbacks.

- ***Technical issues related to the deployment of the new feature***

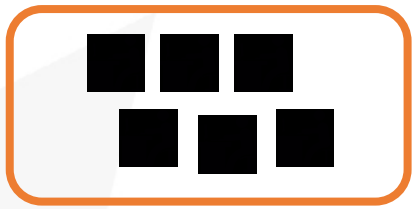
  Some technical problems, due to the implementation of the new architecture, have raised difficulties in the process of the collection of data.
  This experimental phase helped a lot in the identification of possible errors or bugs.

DATEV

Participation to this testing phase

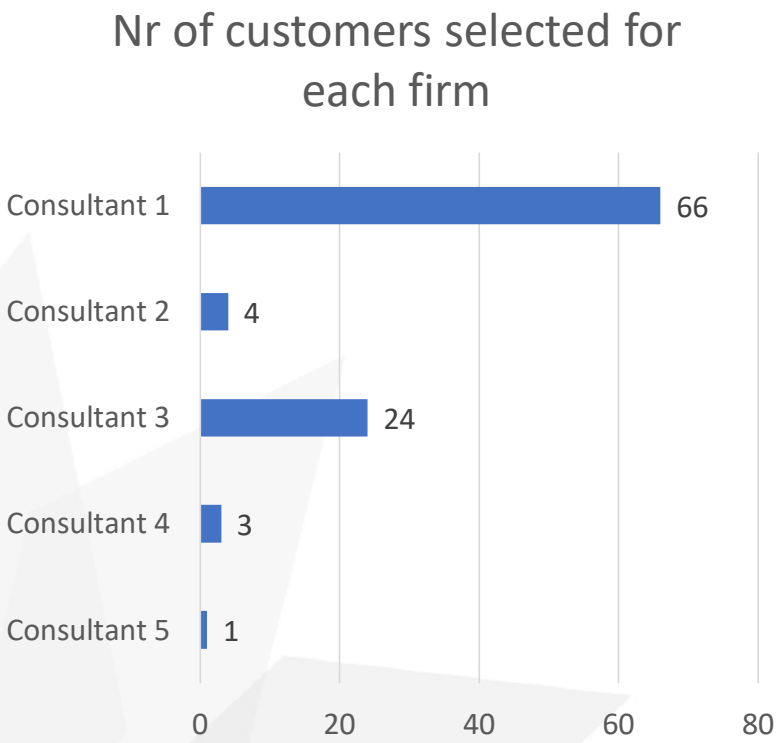


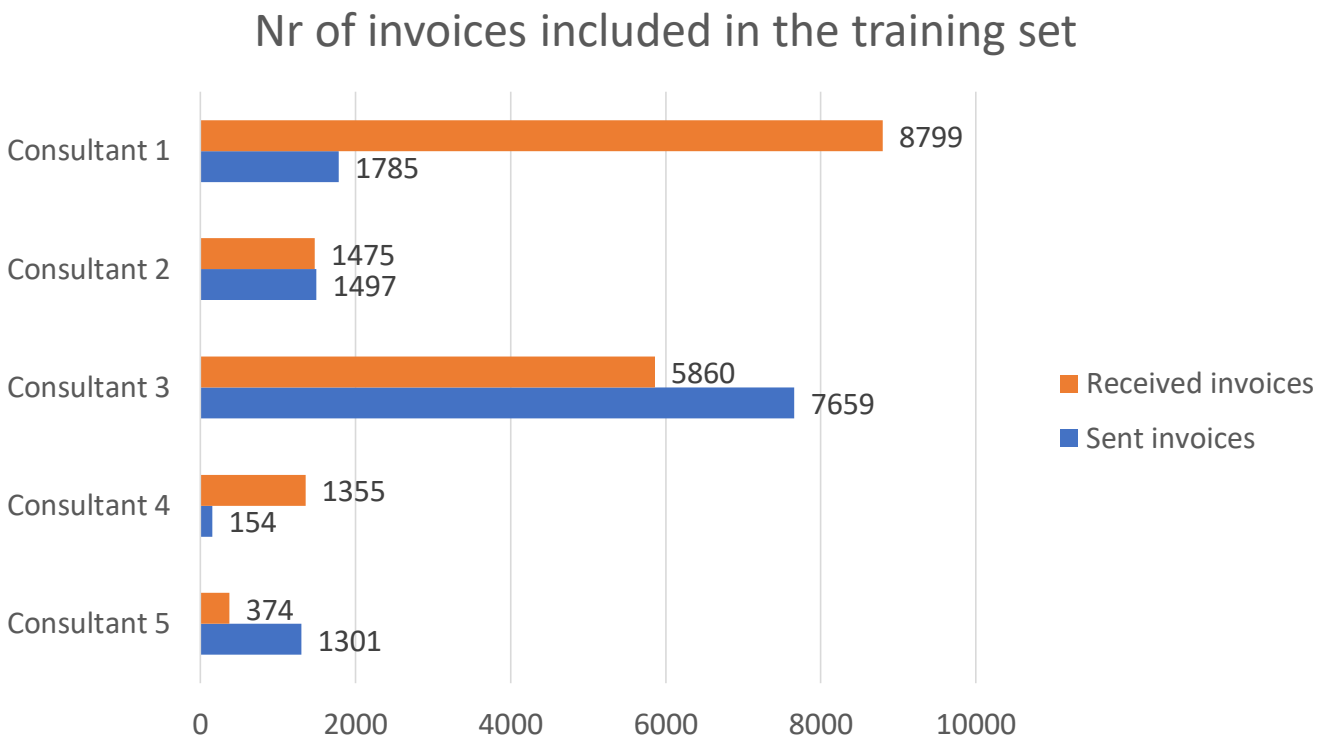
5 tax consultants are actively testing the procedure and sending feedbacks



6 tax consultants expressed interest in the experimentation and are waiting to be enrolled

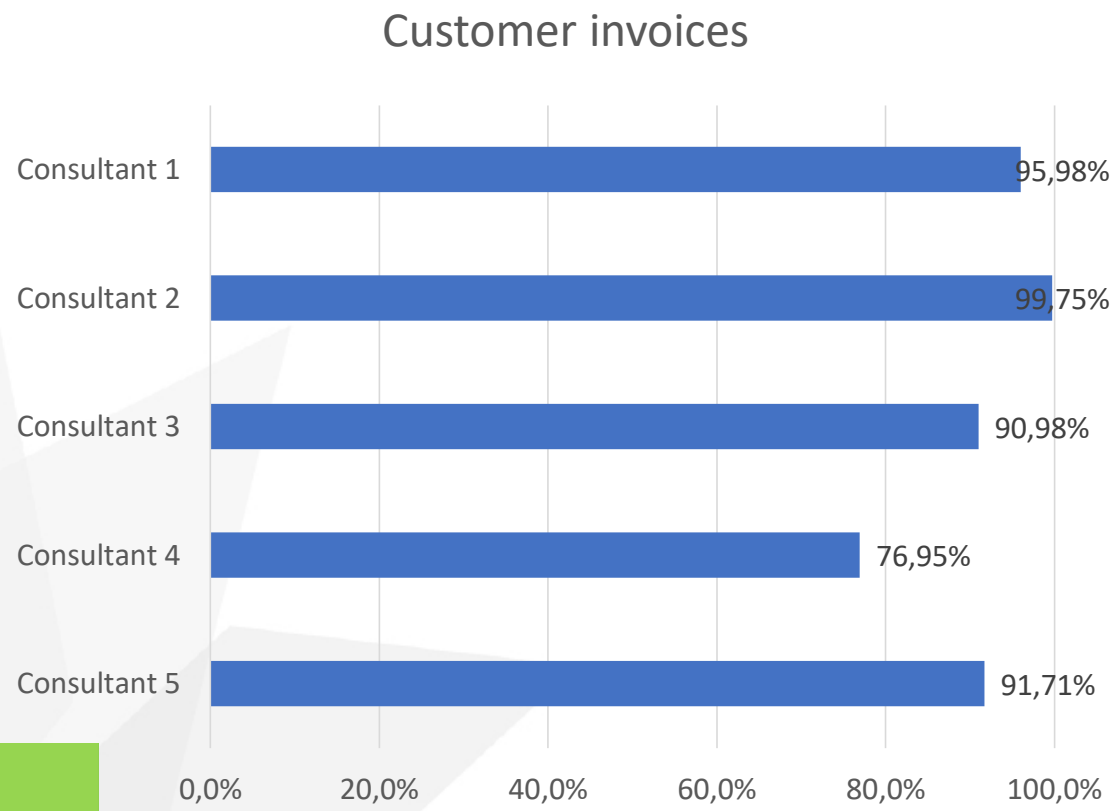
Description of the data

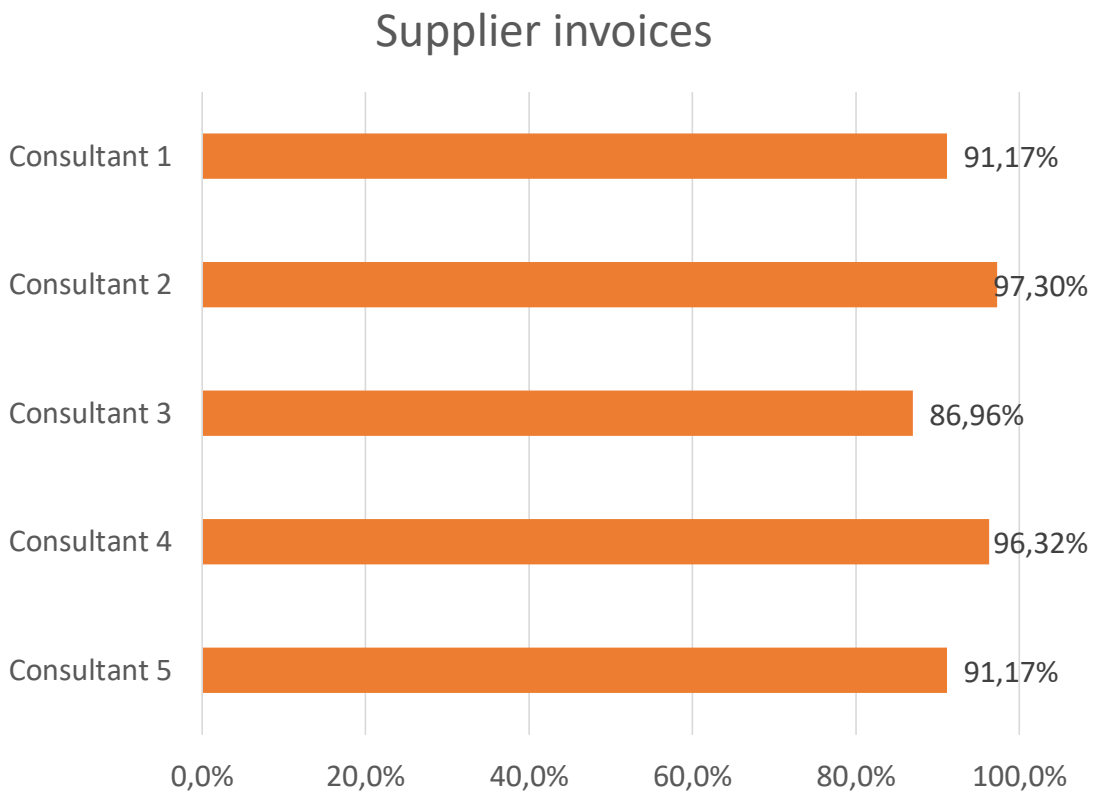




Accuracy for the account prediction

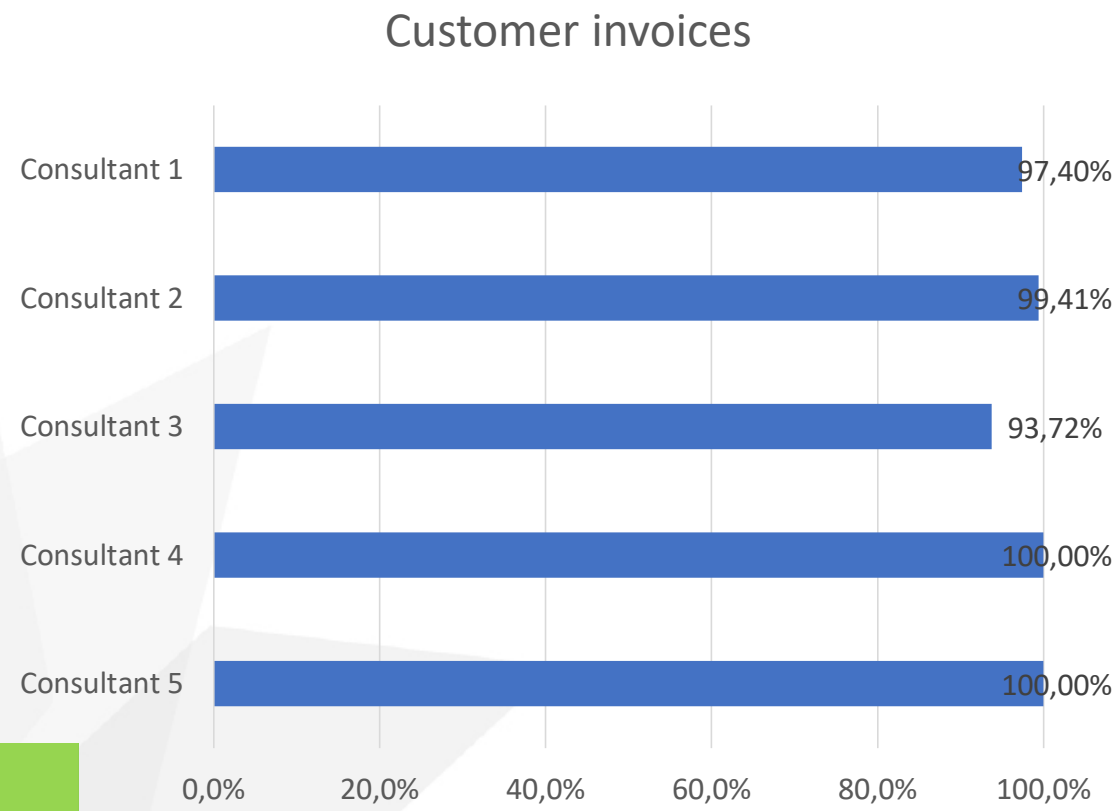
Accuracy computed using 10-fold cross validation.

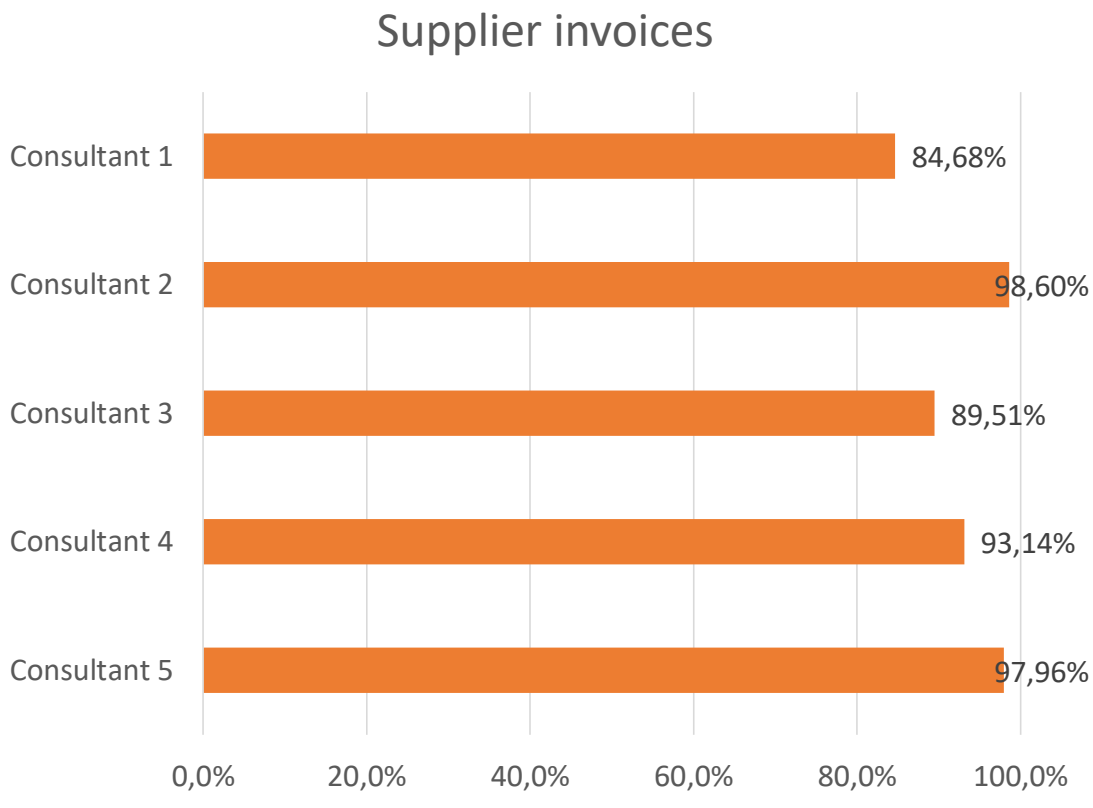




Accuracy for the VAT codes prediction

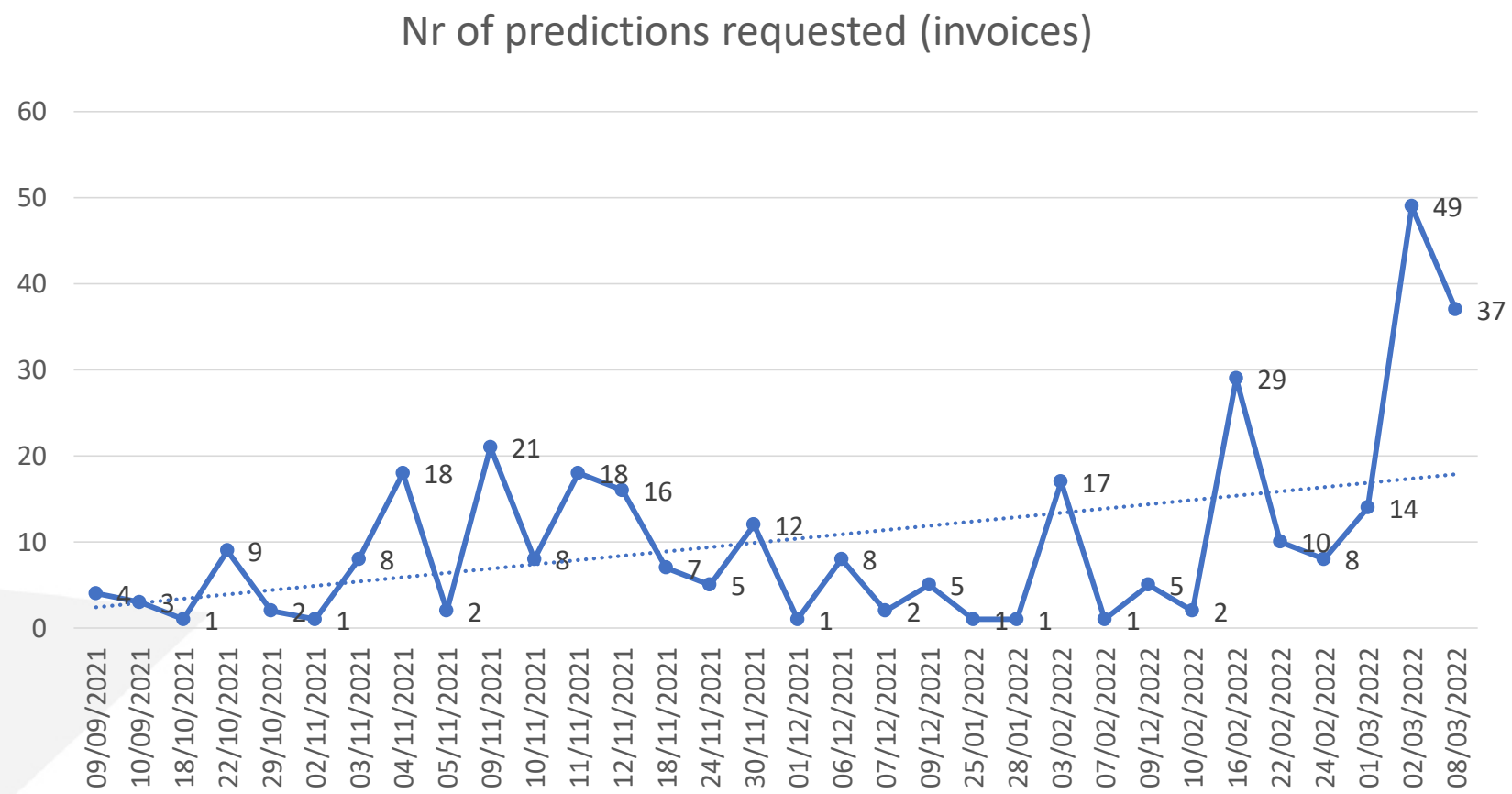
Accuracy computed using 10-fold cross validation.



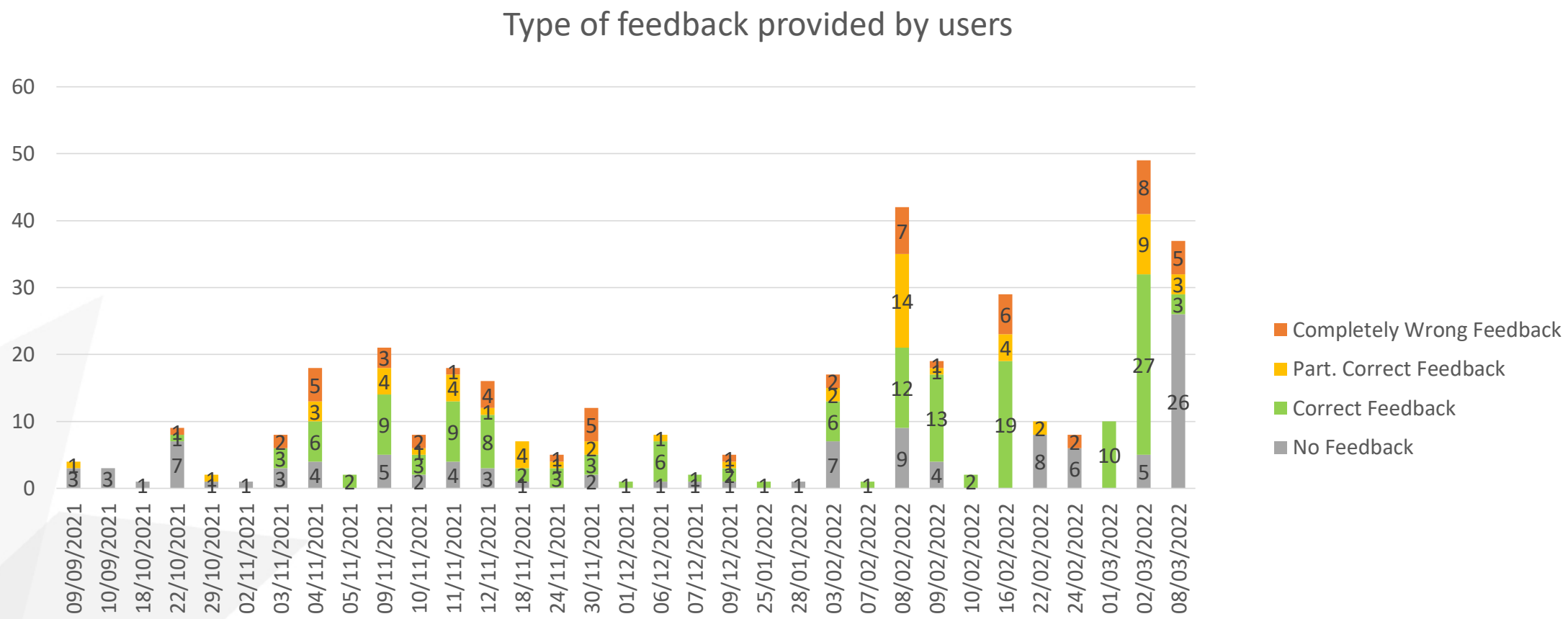


Invoices processed

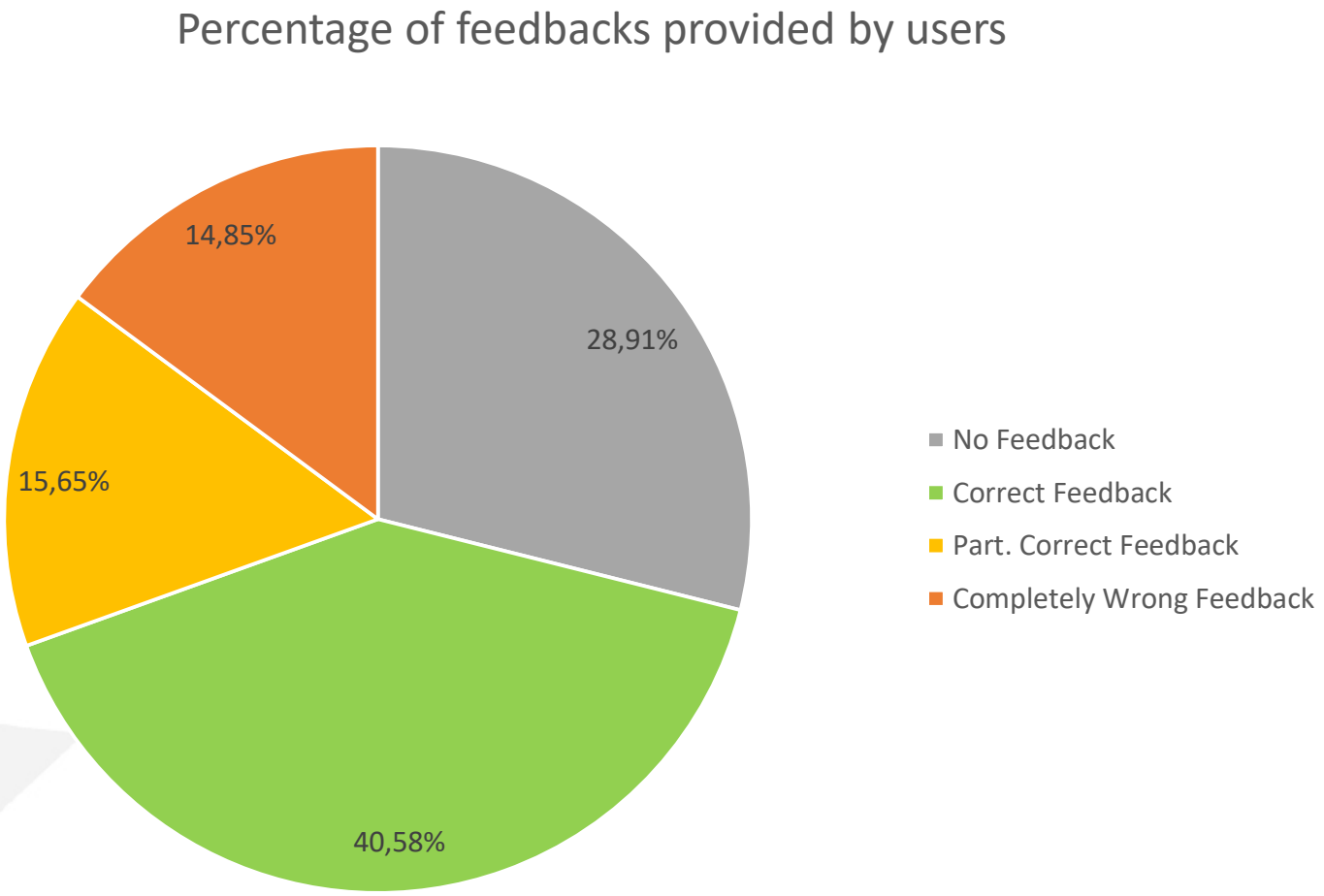
325 invoices sent by users to verify the proposals of the algorithm



Feedback data



Feedback data - percentage



Conclusions and Q&A

DATEV

Conclusions

Phase 1 is still ongoing to collect statistics on different databases of tax consultants.

Future steps to make the experimental phase more concrete:

- Include other tax consultants
- Integrate the proposals of the ML algorithm in the creation of bookkeeping journal entry

A lot of things to do and to improve…

- Exploit the hierarchical structure of codes
- Examine other nlp models
- Improve the matching algorithm
- …

DATEV

Some final comments

- Collaboration with the experts of accounting in Datev has been crucial for the first phase of the analysis, to understand the domain of the problem.

- A data science project is not only about the implementation of a Machine Learning algorithm.

- Design an efficient and usable architecture is still an open problem for us and the most challenging issue.

What about my experience?
- possibility to combine my research activities and the real need of a software company.
- learn how to transform a research project into a usable feature of a software.

Why choose an industrial PhD?
- It depends on your interests and your background
- Start from a real application can be more stimulating and interesting

Thank you for your attention!

Any questions or comments?

DATEV