



SAS Viya – Offenheit für unterschiedliche Benutzertypen

Am Beispiel der Analyse der Forecast Errors

Franz Helmreich

Gernot Engel

Matthias Svolba

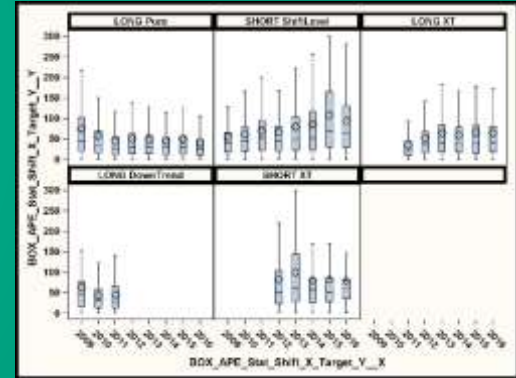
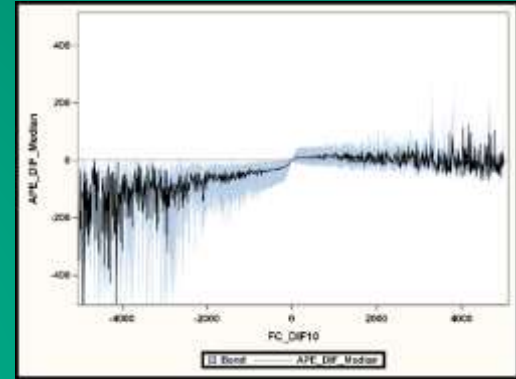
Gerhard Svolba

Data Science in Action: #4

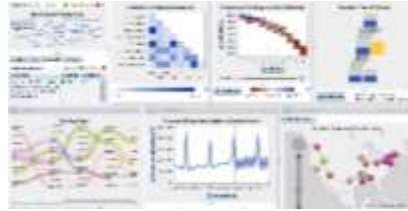
Explaining Forecast Errors and Deviations

*Do the demand planners really improve
forecast accuracy with their manual
overwrites?*

Linear Regression
Quantile Regression
Descriptive Statistics



Intention dieses Beispiel: Demonstration der Offenheit der SAS Analytic Plattform für unterschiedliche Benutzertypen



Office
Integration



One Integrated Solution for Different User Types



„Was bisher geschah“ (1)

Fachliche Fragestellung im Unternehmen

- Monatliche Nachfrage wird mittels Zeitreihenmodellen vorhergesagt.
- Produkte mit langer und kurzer Absatzhistorie (langjährige Artikel sowie Fashion)
- Analyse des Vorhersage-Fehlers
 - Welcher Vorhersage-Fehler wird von 50%, 75% meiner Vorhersagen nicht überschritten?
 - Welchen Faktoren führen zu geringen Vorhersagefehlern?
 - Gibt es zeitliche Trends und saisonale Muster im Vorhersagefehler?
- Faktoren
 - Stammdaten: PRODUCT_AGE, PRICE_INDEX, LAUNCH_MONTH, PRODUCT_GROUP
 - Stat. Forecasting: MODEL, LEAD_TIME, TARGET_YEAR, TARGET_CALENDAR_MONTH

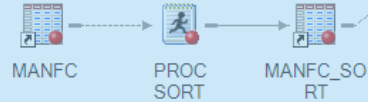
„Was bisher geschah“ (2)

Datenaufbereitung

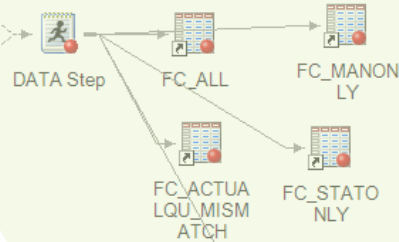
Statistischer Forecast



Manueller Forecast



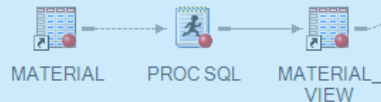
Datenqualitätsprüfung



Berechnung abgeleiteter Variablen



Artikel-Stammdaten

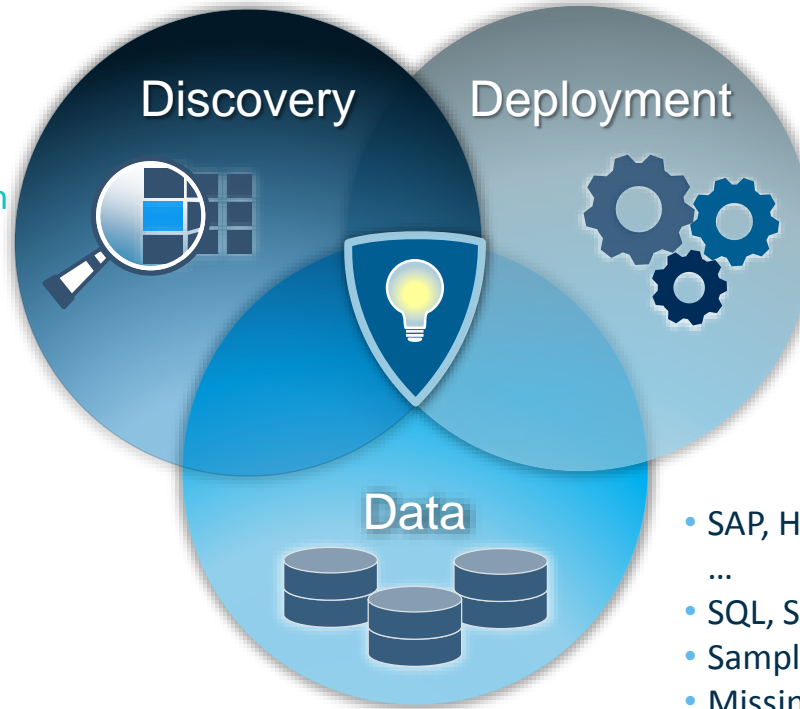


$$\text{APE_STAT} = \frac{\text{abs}(\text{statfc} - \text{actual})}{\text{actual}} * 100$$

→ Absolute Percentage Error

Data Mining und Machine Learning mit der SAS Analytic Plattform

- Logistic Regression
- Linear Regression
- Generalized Linear Models
- Nonlinear Regression
- Ordinary Least Squares Regression
- Decision Trees
- Partial Least Squares Regression
- Quantile Regression
- K-means and K-modes Clustering
- Principal Component Analysis
- Random Forest
- Gradient Boosting
- Neural Networks
- Support Vector Machines
- Factorization Machines
- Network Analytics/Community Detection
- Text Mining
- Boolean Rules
- Auto-tuned Hyper-parameters



- Assess Supervised Models
- Modellverwaltung
- Deployment
- Laufende Validierung
- Modell-Retirement
- Retraining

- SAP, Hadoop, Streaming, rel.DB, ...
- SQL, SAS Dastep, Matrix
- Sampling and Partitioning
- Missing Value Imputation
- Variable Binning
- Variable Selection
- Transpose

4 Benutzerrollen analysieren die Fragestellung

One Integrated Solution for Different User Types



Business Analyst

Gerhard Svolba
SAS Visual Analytics



New-to-SAS Statistician

Gernot Engel
SAS Studio Tasks



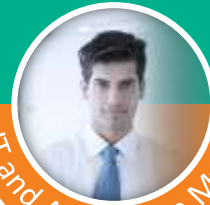
SAS Data Scientist

Franz Helmreich
SAS Studio Program



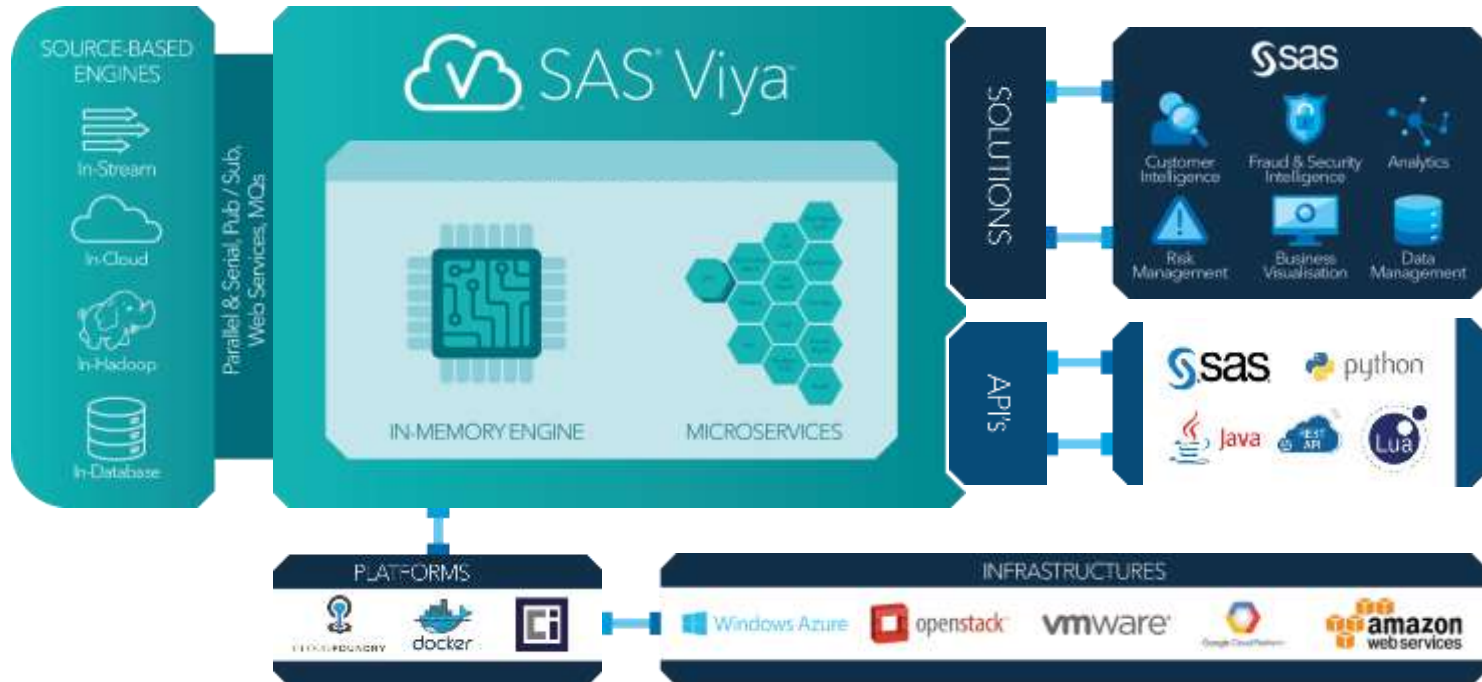
Open Source Data Scientist

Matthias Svolba
Python (Jupyter)



IT and Application Mngt.

Überblick über die SAS Analytic Plattform



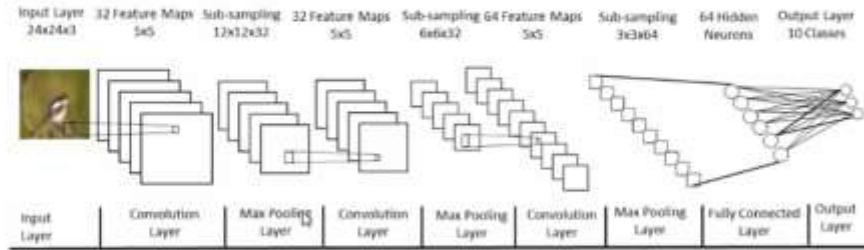
Opening the SAS Analytic Platform via Different Interfaces



Start the SAS Cloud Analytic Server
Load the DeepLearning Action Set

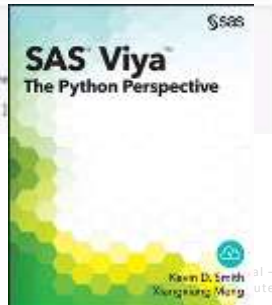
```
jupyter c1b10_tech_exchng_demo Last Checkpoint: 30 minutes ago (unsaved changes)
Python
In [ ]: s = sas.CAS('cas01.usn.eds.com', 11775)
s.sessionprop.setsessionid('CASUSER', timeout=3.1536E7)
s.loadactionset('deeplearn')
```

Define the Network Architecture

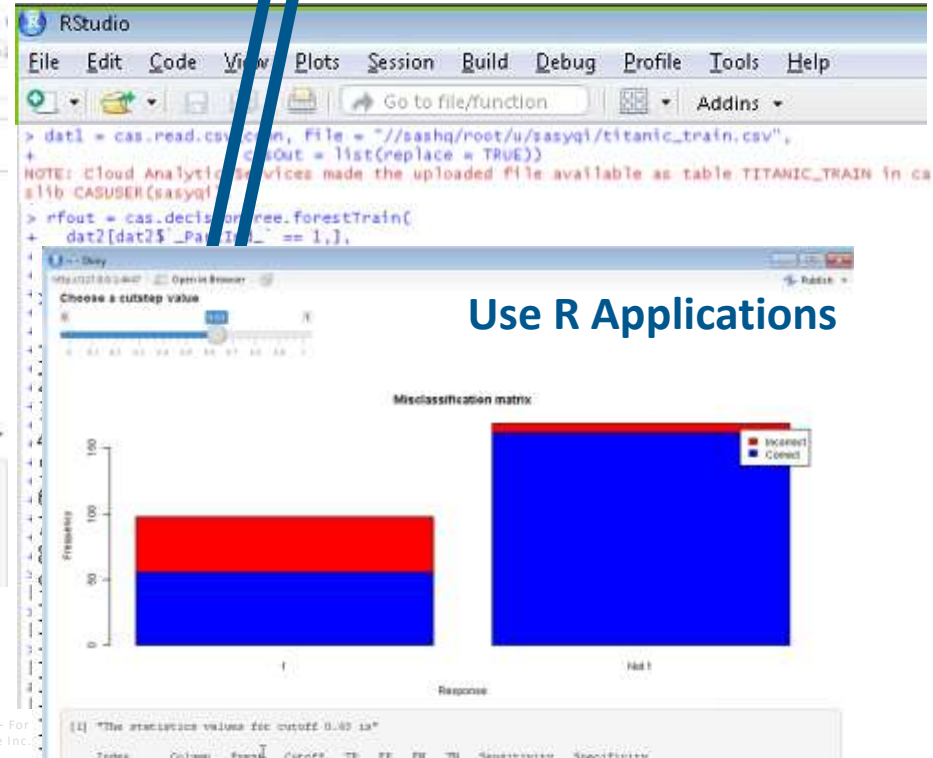


```
In [ ]: s.createModel(model=dict(name='convnet', replace=True), type='CNN')
s.addLayer(model='convnet', name='data', type='input',
            inputOpts=dict(channels=3, width=24, height=24, scale=1))
s.addLayer(model='convnet', name='conv1', type='convolution',
            convOpts=dict(nfilters=32, width=3, height=3, stride=1, init='nrg2'), srcLayers=['data'])
s.addLayer(model='convnet', name='pool1', type='pooling',
            poolingOpts=dict(width=2, height=2, pool='max'), srcLayers=['conv1'])
s.addLayer(model='convnet', name='conv2', type='convolution',
```

Define the Network Layers



Use the CAS Random Forest
Display the Results in R-Studio



Key Takeaways

Analytics und Data Science sind da um Ihnen zu helfen!

- Sie sehen ein klareres, objektiveres Bild Ihrer Daten und Analyse-Subjekte
- Sie erhalten explizite Ergebnisse anstatt die Nadel im Heuhaufen zu suchen
- Die Daten sprechen zu Ihnen und Sie erhalten die Ergebnisse automatisch statt manuell
- Do it again! – Behandeln Sie Ihre Modelle als “Asset” und wiederholen Sie Ihre Analyse

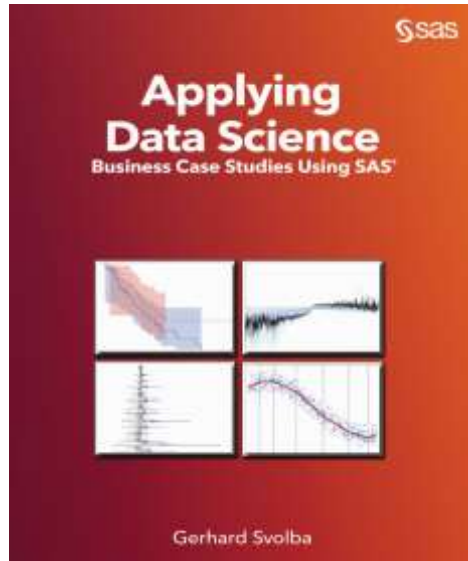
Machine Learning and Data Science sind das Kernstück der SAS Analytic Platform

- Umfassendes Set an Methoden – Entdecken und Produktivstellen
- Offen für unterschiedliche Benutzertypen (Coding, Point&Click, SAS, R, Python, ...)

More Information

Gerhard Svolba – Principal Analytic Solutions Architect

sastools.by.gerhard@gmx.net



- Applying Data Science – Business Case Studies Using SAS, SAS Press 2017
- Eight Case Studies showing how Data Science and Analytics can be applied to provide insight into your data and improve your business decisions
- [http://www.sascommunity.org/wiki/Applying_Data_Science - Business Case Studies Using SAS](http://www.sascommunity.org/wiki/Applying_Data_Science_-_Business_Case_Studies_Using_SAS)