No-Code AI for Citizen Data Scientists

Squirro’s AI Studio

Make AI usable for everybody through Augmented Intelligence
WHO AM I?

Speaker’s Introduction

Technologist and entrepreneur with a passion for building applications that push the limits. Also a father, runner, cyclist and language nerd.

Patrice Neff
Co-Founder of Squirro
Meet Squirro

Global Company
Augmented Intelligence Software Provider
Built by team of serial entrepreneurs
Global presence and partner network

A different approach to AI
A powerful Insights Engine for Decision Intelligence
Tangible, industry and issue specific AI-Apps
Build your own apps with the Data Science Studio

Reference Customers

Investors

© 2021 Squirro AG
most valuable

actually used
Insights:

80% of all data is locked away. Reason: The data is unstructured, siloed and fast changing.

Yet, hidden in it are insights into your business, your clients, your competitors & markets.

Apply these insights to your business and get a direct positive top- and bottom-line impact.

Source: Hopkins et al. Forrester
Our Mission:

Direct top / bottom line business impact

Ready to use, pre-configured AI-Apps

Industry and issue specific pre-trained AI-Models

Easy to deploy: Cloud, private cloud or on-premise

<table>
<thead>
<tr>
<th>Insights</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales Insights</td>
<td>Opportunity Sourcing: Reach revenue targets and avoid client churn; better opportunity prioritization; + 5% revenue</td>
</tr>
<tr>
<td>Marketing Insights</td>
<td>Stay Ahead: See trends earlier, get a better understanding of markets and competition; 22% better coverage</td>
</tr>
<tr>
<td>Service Insights</td>
<td>Resolution Automation: Delight your clients; Auto-classify cases, automate resolution; cut 30% mean-time-to-resolution</td>
</tr>
<tr>
<td>Risk Insights</td>
<td>Early Detection: Keep your business safe — spot non-numeric threats, risk deflection; cut 15% effort</td>
</tr>
<tr>
<td>Cognitive Search</td>
<td>Insights Discovery: Unified search across all data stored in any platform; easy insights discovery; +32% data ROI</td>
</tr>
</tbody>
</table>
Squirro’s Insight Engine

Application Layer

Gather
Load Enrich Relate
Data Loader Enrich Pipelets Relate Pipelets

Understand
Discover Classify Predict
Auto Discovery Signal Detection Anomaly & Trend Detection

Act
Recommend Automate Visualize
Recommendation Automation Engine Dashboard Builder

Data Science Studio & Machine Learning Service

Cloud / On-Premise Multi-Instance / Scalable Admin / Security

© 2021 Squirro AG
https://info.squirro.com/technology
Best fit for any Situation

Enterprise learning cycle

Little data
No models

Ample data
Strong models

From simpler to more complex methods

Ad-hoc search
Smart Filters
Naive Bayes
SVM
Random Forest
Deep Learning

Squirro Machine Learning Service

Predict
Classify
Recommend
Trigger Automation
Engine action
So: immediate or accurate?

Model A

More accurate long term
Slower learner

Model B

Less accurate long term
Faster learner
Flexibly choose the best Option

Accuracy

Model B  Time  Model A
Squirro AI Studio

Challenge

• Lacking quality & representative training data

• Limited knowledge of pattern analysis, model training, optimization, and deployment

Solution

• The Squirro AI Studio supports the full ML lifecycle

• Self service use for any citizen data scientist

Benefits

• ML made simple for non-data scientists

• Speed up creation of models by 78%
Welcome to the Squirro AI Studio

Follow the steps below to build and deploy your models from scratch

1. **Candidate Sets**
   Start by creating a Candidate Set of documents you would like to label. Each candidate set consists of a shortlist of documents.

2. **Ground Truth**
   Use the candidate set from the previous step to create Ground Truths. Create labelled data sets and proximity search rules to be fed to the model.

3. **Models**
   Create Models by combining a Ground Truth with a ready to use ML template.

4. **Validation**
   Validate the models to choose the best performing one. Deploy your model to the Squirro platform where you can start using it within a pipeline.

5. **Publish**
   Publish your ready to use models so they can be added to the pipeline.

---

Start by creating a Candidate Set for the documents you would like to label.

Each candidate set consists of a shortlist of documents.

Use a candidate set to label data and create a Ground Truth set.

Create labelled data sets and proximity search rules in a WYSIWYG interface.

Create a model by combining a Ground Truth set with your ML template of choice.

Choose from Proximity Search, Naïve Bayes, SVM, Random Forrest, Deep Learning, etc.

Validate the accuracy of your model to choose the best performing one.

Adjust the model based on user and usage feedback.

Publish your ready to use models.

Deploy your model to the Squirro platform where you can start using it within a pipeline workflow.

---

For a short demo video see here: http://go.squirro.com/ais
For full documentation see here: http://squirro.com/ai-studio
The first step for creating a data classification model is to select data which is most suitable for labelling.

As often high-quality training data is not a given, this step in selecting good quality data for the labelling is essential.

By defining search queries new matching data will be added automatically to the candidate set.

Filter
Use the full power of Squirro to quickly identify good candidates

Select
Select the best candidates from the data available

Save
Create iterations of your candidate set
Ground Truth Definition

Once the best candidates for labelling are identified, the data must be labelled to define the ground truth for the model.

Depending on the target classification model, data can be labelled either on a document or a sentence level.

While for document-level models, the whole document is labelled, sentence-level tagging enables the labelling of individual sentences within a document.

---

Label Data

- Four approaches to quickly label data
- Labelling on document or sentence level

Labelling

- System pre-labels based on input and Auto-ML*
- Label each item & ability to create rule-based models for faster results

* Roadmap 2021; For full documentation see here: http://squirro.com/ai-studio
Model Creation

A Ground Truths set may be combined with a pre-defined Machine Learning Template to create a model.

The ML templates comprise steps such as language handling, filtering, data cleansing, tokenization, etc.

The DSS enables a fast and efficient testing of different NLP algorithms and retraining with updated datasets.

Build Models
A model is built by choosing a Ground Truth, an Encoder, and an Algorithm.
To simplify we allow to select from pre-configured ML templates.

Setup
Further configure the model per your requirement.

For a short demo video see here: http://go.squirro.com/ais
For full documentation see here: http://squirro.com/ai-studio
Model Validation

Trained models may be validated on the training dataset and compared to previous iterations & configurations.

A solid validation process covering accuracy, precision, & recall ratios is crucial for quality classification results.

Besides validating a model, you can also compare model performance to other models built with different ML templates.

Validate Models
For any model created run a full validation process to determine model accuracy and quality.

Compare Models
Compare the performance of the models to previous models.
Model Publishing

Models that have been successfully validated can be published to Squirro and added to the classification pipeline.

Existing classification pipelines can be updated with new and / or retrained models.

You may classify your own data in your own setup, too; open an API end-point for on-demand data classification.

Publish Models

Once ready, publish models to the Squirro platform or update models on live projects.

API based classification

Open an API end-point to make data classification available to your own setup.
ML Ops Framework

Release Cycle
- Integration Testing
- Operational Validation

Development Cycle
- Model Building
- Model Validation
- New Data & Feedback
- Business Demand
- Data Source Discovery

Operations Cycle
- Model KPI Validation
- Model Activation
- Model Deployment
- Production Audit
- Model Behavior Tracking & Feedback

Mgmt & Governance:
Deployment, Explaining, Complying, Feedback
ML Ops Squirro Setup

For a short demo video see here: http://go.squirro.com/ais
For full documentation see here: http://squirro.com/ai-studio
Thank You

Patrice Neff
Co-Founder

patrice@squirro.com