Ablauf des Learners für den Prototyp

### Erneutes Lernen

Input Daten:

Der Learner erhält im Produktivsystem folgende Inputdaten als CSV- Datei zum weiteren Lernen:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| learnerId, | userId, | documentId, | message, | documentAffiliation, | rate, | favor, |

Datentransformation

Die Input Daten werden mit den Modulen *Edit Metadata*, *Clean Missing Data* und *Select Column in Dataset* für das Lernen aufbereitet.

*Edit Metadata* wandelt die userId in einen String um, da sonst der Learner die userID als numerischen Wert verarbeitet, was den Output des Learners beinträchtigen würde.

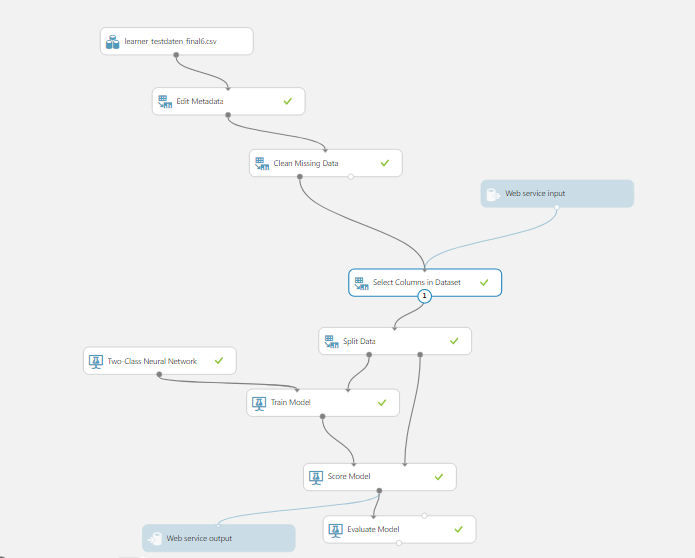
*Clean Missing Data* wird gebraucht, da in der favor-Spalte, bei einer Favorisierung des Dokuments vom User, *true* angezeigt wird. Bei einer Nicht-Favorisierung bleibt die Spalte jedoch leer. Wenn die Spalte aber leer bleibt, funktioniert der Learner fehlerhaft. Um dies zu umgehen, werden durch *Clean Missing Data* die leeren Zeilen durch den Wert *false* ersetzt.

*Select Column in Dataset* filtert die Inputdaten, sodass lediglich userId, documentId und favor selektiert sind.

Modell trainieren

Um das Modell zu trainieren, werden Daten mit einem Split Data Modul im Verhältnis 80/20 aufgeteilt. 80% der Testdaten fungierten dabei als Trainingsdaten für den Learner. 20% als Abgleichdaten des Gelernten. Für die Umsetzung des Learners wurde, wie bereits erwähnt, ein *Two-class neuronal network* Modul verwendet.

Folgende Abbildung veranschaulicht das Experiment-Modell des Protoyp Learners.



*Abbildung 2: Experiment-Modell des Learners für den Prototyp*

Das Experiment-Modell des Prototyp Learners ist abrufbar unter:

<https://gallery.cortanaintelligence.com/Experiment/Final-Learner-Dokument-Nutzer-Final-Parameter2>

### Abrufen des Predictive Modells

Das Predictive Modell benötigt als Input:

* userId
* DocumentId

Ein Beispiel des Outputs des Learners im Predictive Modell sieht wie folgt aus:

["100213199594809398744","1SAI8Y7j6kNfbbdYVaUpEPipzy-WhcvmuPRynSg76UIM","True","0.938885152339935"]

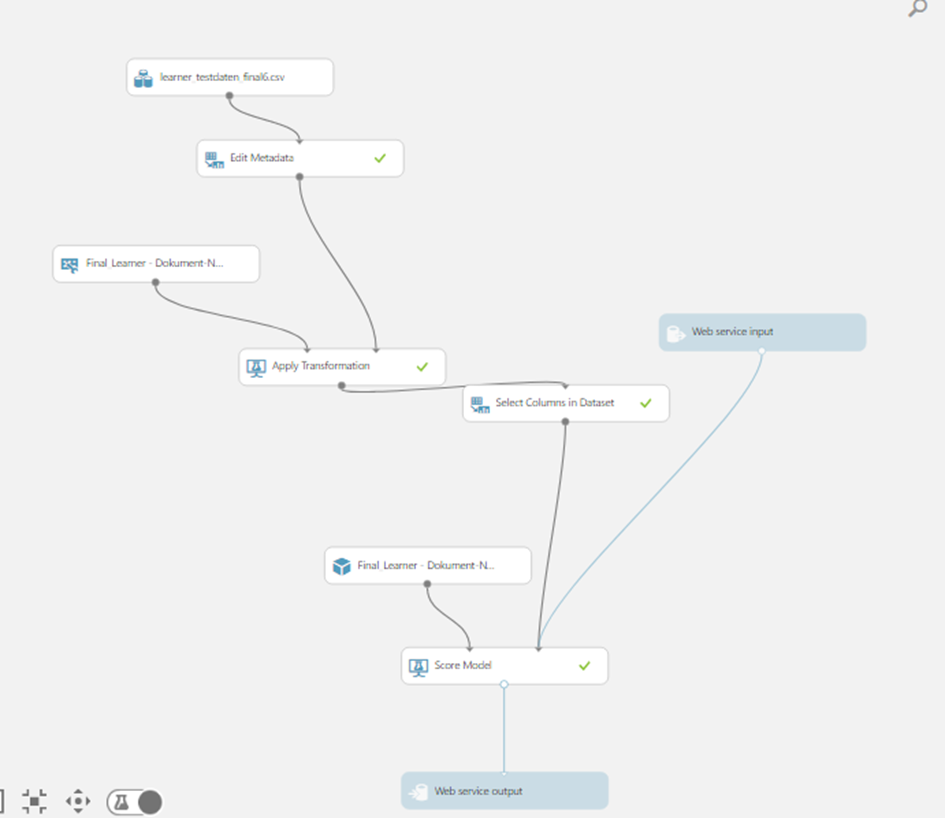
Der Output besteht dabei aus:

* userid,
* documentId,
* Favorisierung (true/false)
* Score Probability Wert des Learners

Dieser wird in JSON ausgegeben.

Das Predictive-Modell des Prototyp Learners ist abrufbar unter:

<https://gallery.cortanaintelligence.com/Experiment/Final-Learner-Dokument-Nutzer-Final-Parameter2-Predictive-Exp>



*Abbildung 3: Predictive-Modell des Prototyp Learners*

### Testcase des Prototyps

Anhand der für den Prototyp angelegten User und Dokumente wurde ein Testcase für den Prototyp erstellt.

Die GUI-Gruppe hat folgende User im Prototyp System angelegt:

|  |  |  |  |
| --- | --- | --- | --- |
| Vorname | Nachname | Mail-Adresse | userID |
| Lisa | Meier | lisameier3141@gmail.com | 100213199594809000000 |
| Haruki | Hakashima | harukihakashima@gmail.com | 109237104808355000000 |
| Hang | Outs | hangouts767@gmail.com | 102231370704961000000 |
| Claude | Systém | claudesystem@gmail.com | 103097844766993000000 |

*Tabelle 1: Angelegte User des Prototyp Systems*

Die SemRep-Gruppe hat folgende Dokumente im Prototyp-System angelegt:

|  |  |
| --- | --- |
| DocumentID | DocumentName |
| 1Cl1k0OYjdubR0\_BqiFp2xaRGeRpqqUz8pucX7iQwIQ8 | Anforderungen |
| 1kapEE9ui6oYu-uUZF6qjg-j85iJz\_w7lNmtyLHZA7lc | KickOff |
| 1mC30R185Km9Y9HoE2uUgVDJi41IXKelPJEBhBeXH2PY | Aufgabenliste |
| 1jQFZmcS\_\_-CtScjqd3g5KKM8xepPnMSqaepy2ag2jNc | Meilensteine |
| 1gnH6sv3VqGEveSSf0SoS7dJqWaUyUnZm6-QrdiH7vCM | Projektplanung |
| 1Ktol-SWvAh8pnHG2O7HdPrfbEVZWX3Vf2YIPYXA\_8gI | Finanzreport |
| 1jYsTPJH8gaIiATix9x34Ekcj31ifJMkPNb0RmxnuGxs | Kostenaufstellung |
| 1SAI8Y7j6kNfbbdYVaUpEPipzy-WhcvmuPRynSg76UIM | Kostenplanung |

*Tabelle 2: Angelegte Dokumente des Prototyp Systems*

**Testcase für das Prototypsystem**

Mitarbeiter im Rechnungswesen arbeiten und favorisieren in der Regel folgende Dokumente: Finanzreport, Kostenaufstellung und Kostenplanung. Lisa arbeitet im Rechnungswesen, und hat bisher mit den Dokumenten Finanzreport und Kostenaufstellung zu tun gehabt. Sie öffnet das Dokument Kostenplanung zum ersten Mal. Der Output des Learners gibt preis, dass Lisa mit einer sehr hohen Wahrscheinlichkeit das Dokument Kostenplanung ebenfalls favorisieren wird. Der Ouptut des Learners sieht dabei wie folgt aus:

["100213199594809398744","1SAI8Y7j6kNfbbdYVaUpEPipzy-WhcvmuPRynSg76UIM","True","0.938885152339935"]