

Project proposal

Pavel Semenov

06.01.2023

What is the problem?

I am currently employed at the IBM as a DevOps engineer. I have almost no project management responsibilities, but I am sometimes asked by project manager to give an estimation of time required to complete the task. It is always very difficult for me to give an accurate estimate. Thus, I would like to produce a model, which might be able to approximate the time needed for the task to be finished. I am not sure there is a strong correlation between the task parameters and the resolution time, but I am willing to discover it if it exists. We'll see.

What is my dataset?

I have a collection of all resolved tasks from one of the IBM projects, which have non-zero time spent parameter. I have exported it as an XML with **28791** work items. Each work item has the following parameters:

Name	Description	Example
timeSpent	The time spent on the task in milliseconds	144000000
summary	Short summary of the work item	FST-4320 - Falscher Ablauf beim Versand XJustiz Nachricht nachricht.vag.fehler Version 3.3.x mit dem gültigem Signaturstatus, mit fehlerhafte Schematronprüfung, mit erfolgreiche Plausibilitätsprüfung

description	Multiline description of the work item	<p>Falscher Ablauf beim Versand XJustiz Nachricht <code>nachricht.vag.fehler</code> Version 3.3.x mit dem gültigem Signaturstatus, mit fehlerhafte Schematronprüfung, mit erfolgreiche Plausibilitätsprüfung</p> <p>Es existiert ein FAM VA Verfahren mit Auskunftersuchen. Eine XJustiz Nachricht <code>nachricht.vag.fehler</code> Version 3.3.1 mit Signaturstatus unknown, mit fehlerhafte Schematronprüfung und erfolgreiche Plausibilitätsprüfung wurde an Postfach gesendet. Es läuft der AF002 Normalablauf (Schritte 1-9). Im Schritt 10. stellt System fest, dass Regelverstöße bei der Schematronprüfung vorliegen.</p> <p>FEHLER IST: Aus dem Verfahren wurde die XJustiz Nachricht <code>nachricht.vag.fehler</code> gesendet und die eingehende XJustiz Nachricht <code>nachricht.vag.fehler</code> landete im Verfahren. SOLL: Es läuft der AF002 Normalablauf (Schritte 1-9). Im Schritt 10. stellt System fest, dass Regelverstöße bei der Schematronprüfung vorliegen und läuft den alternativen Ablauf AF10a erfolgreich weiter. Es wird eine Rückmeldung, XJustiz Nachricht <code>nachricht.vag.fehler</code> Version 3.3.1 mit Verarbeitungsergebnis: FE00110 Beim Validieren gegen das Schema ist ein Fehler aufgetreten! Details: "Fehlermeldung aus Parser".) erstellt und versendet. Das System führt die Aktivität "067_Nachricht aus Zwischenspeicher löschen" ("Nachrichtentyp") für die eingegangene Nachricht aus. Die eingehende XJustiz Nachricht wird nichts in Verfahren gespeichert, sondern wird gelöscht.</p>
priority	Priority assigned to the work item. Ranged from 1 to 5	Prio 2
severity	The severity of the work item. Either Blocker or Normal	Blocker

category	Category of the work item. Might be Release-Test, KM (KonfigurationManagement), Maintenance, etc. (fixed collection of values)	Release-Test
type	Type of the work item. Internal Defect/Issue/Action/etc. (fixed collection of values)	Issue

What is my plan?

I haven't given this part a lot of thought yet. However, I can think of a few steps. Firstly, I would need use some NLP technique (Word2Vec?) to transform the text into vector. Then, I can use it along with other parameters as an input to my Neural Network. Lastly, I would play with its parameters to get the best results. I would greatly appreciate your feedback on this proposal and plan in particular, because my ML knowledge is not sufficient enough to say if this can be done the proposed way.