

Use of AI and ML for RHD content

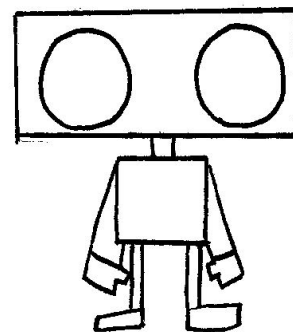
POC for intelligent content search and customer recommendations

Outline

- Machine Learning approach
- Our data
- Technologies
- Results
- Potential use-cases

1. Machine Learning approach

- Given enough data to explore, ML system is able to adapt and decide upon the complex patterns seen before
- A problem with big enough amount of solutions can be used to train the learnable ML system
- The trained ML system can eventually overtake the decision making with accuracy higher than a human observer

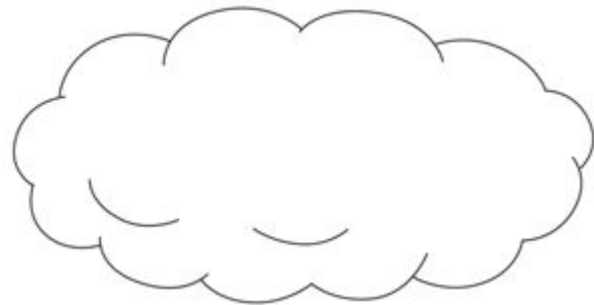


2. Our data: sources

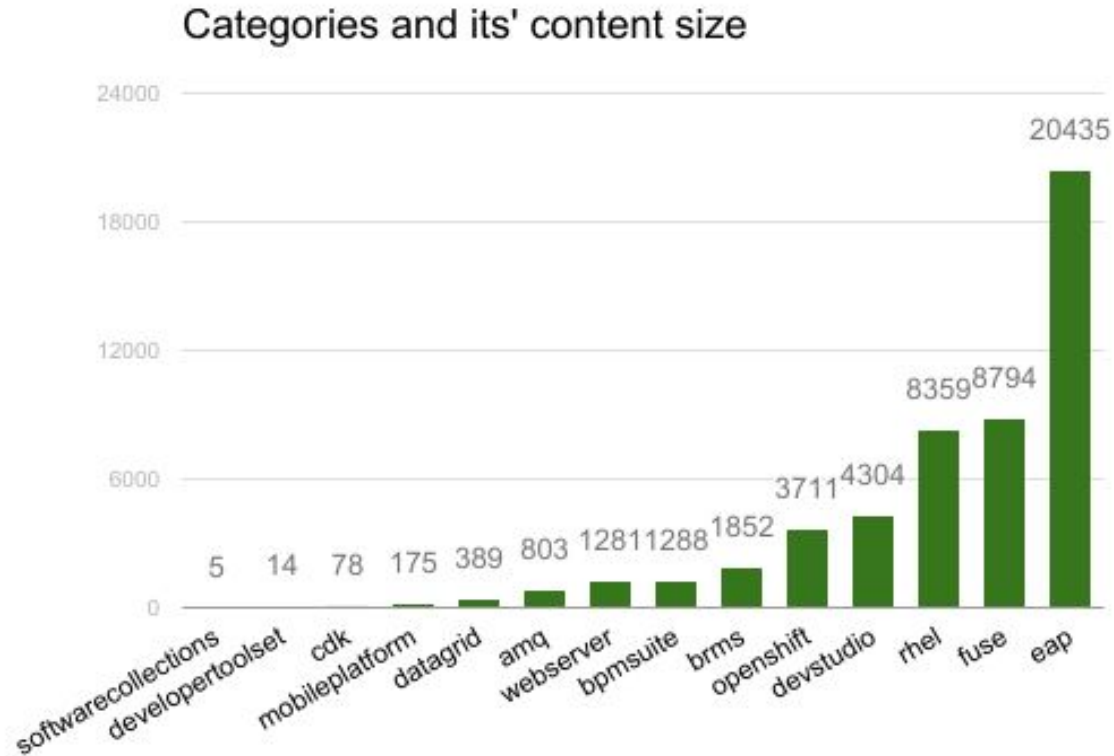
- stackoverflow questions (**13324**)
- access.redhat (**13189**)
- knowledgebase article/solution (**13040**)
- issues.jboss.org (**7965**)
- sbs forums/articles (**3364**)
- ...

total of **51488 documents**

mapped to one of **14 RhT products**



2. Our data: categories



2. Our data

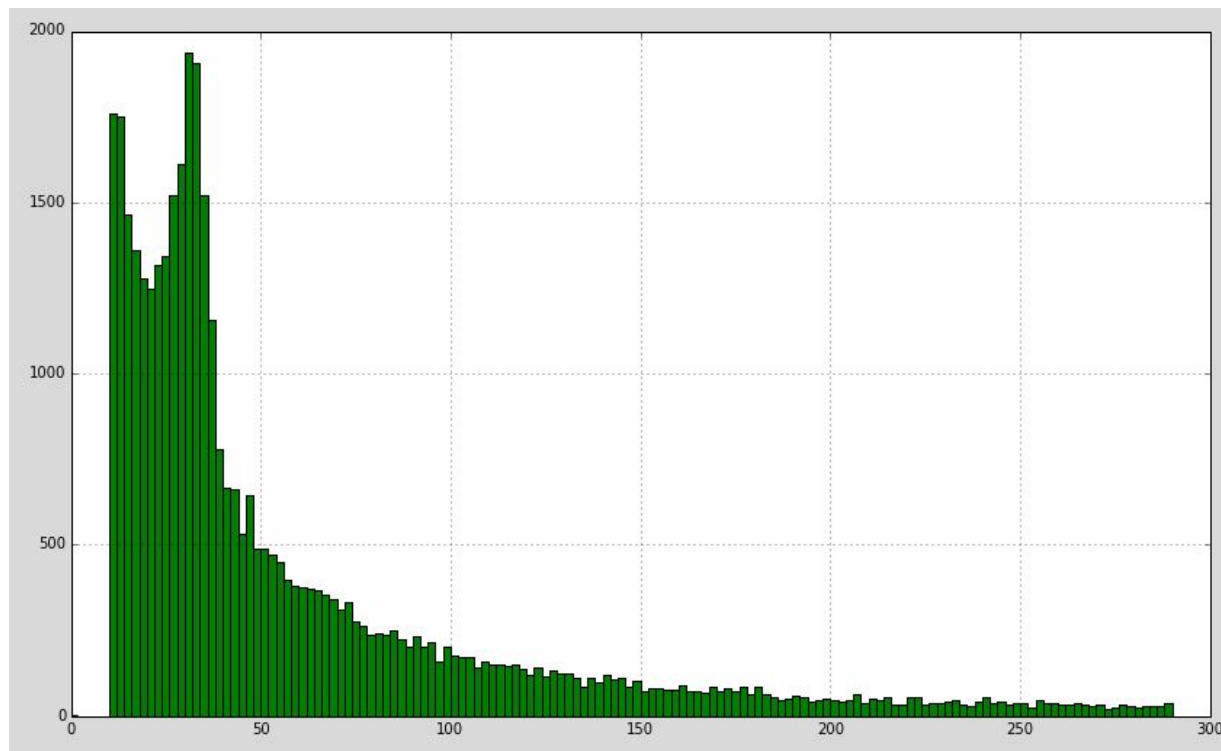
admin shell command ships esb support connect
command idea allow script number commands
containers root vagrant centos bin admin available
commands change rmi registry port changes rmi
registry port management layer existing container
instance change rmi server port changes rmi server
port management layer existing container instance
change ssh port changes secure shell port existing
container instance create creates new container
instance destroy destroys existing container instance
list lists existing container instances start starts
existing container instance stop stops existing
container instance type command help help specified
command

Figure 1: preprocessed document of close-to average size, origin category: fuse

jboss dev studio jboss seam final richfaces want use
richfaces components business logic single jar
component tool file contain suppose customized data
table called application passing parameters sortable
skin want work kind api applications includes possible
ejb seam project kindly provide reference

Figure 2: preprocessed document, origin category: developer studio

2 Our data: categories



Preprocessed documents distribution by size in no. of tokens

3.1 Technologies: What we did with that

1. Word2Vec and Doc2Vec:

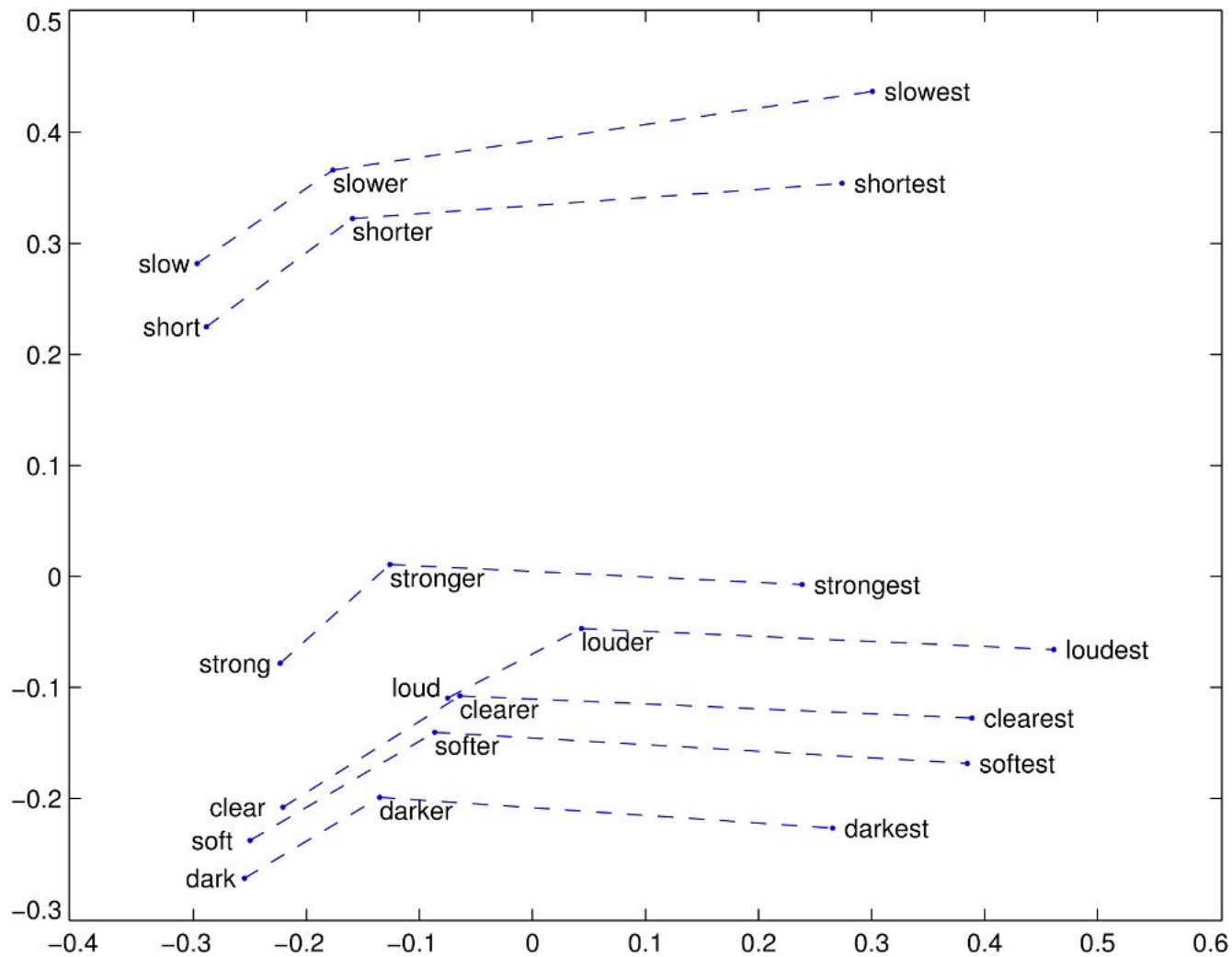
- Technique introduced by Tomas Mikolov from Brno University of Technology (2013)
- Adapted by Google (2014), further optimized in Facebook research (see fastText: 2016)
- Optimized version can be super quickly trained using Neural Network
- There are fancy relations in the created vectors, that ML system can learn

admin shell command ships esb support connect
command idea allow script number commands
containers root vagrant centos bin admin
available commands change rmi registry port
changes rmi registry port management layer
existing container instance change rmi server port
changes rmi server port management layer
existing container instance change ssh port
changes secure shell port existing container
instance create creates new container instance
destroy destroys existing container instance list
lists existing container

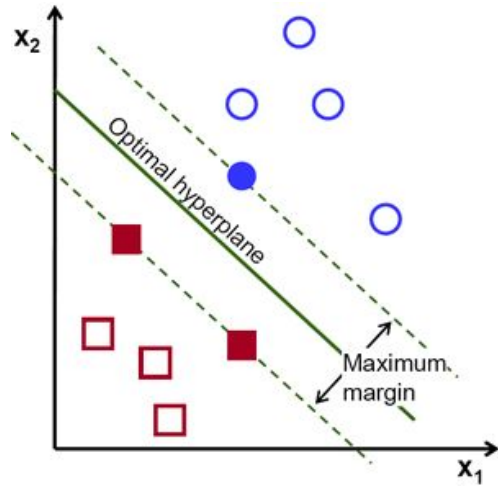


0.52
0.48
0.93
0.01
0.46
0.55
0.64
...
0.77
0.98
0.41

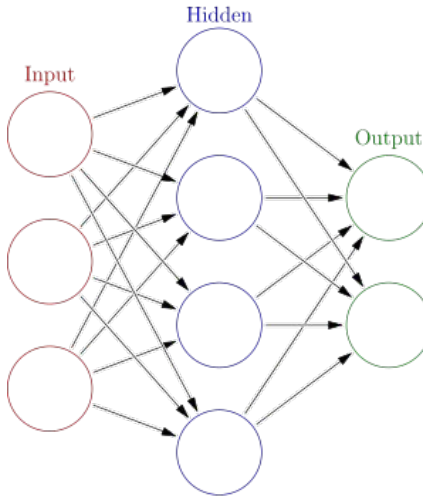
[illegible]



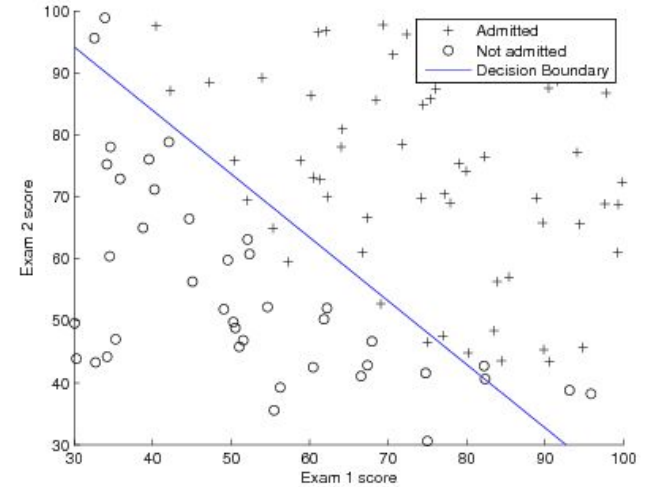
3.2 Technologies: Classifiers



1. Support Vector Machines

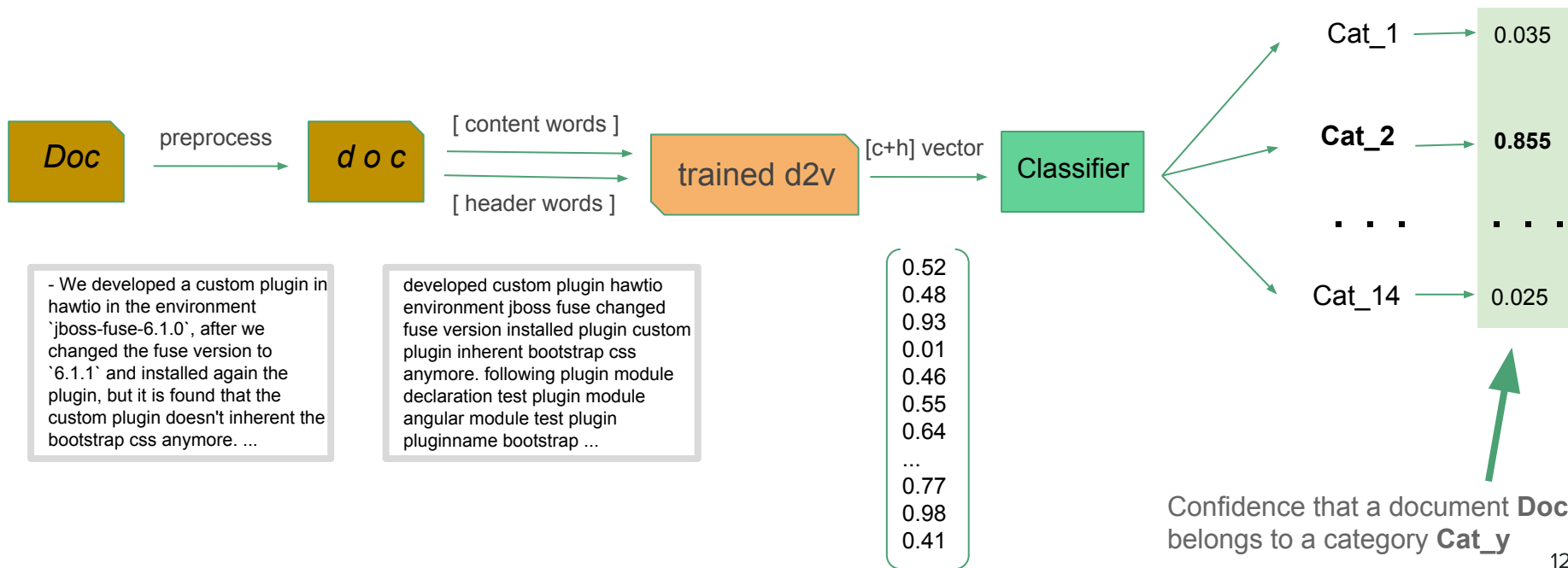


2. Neural Network

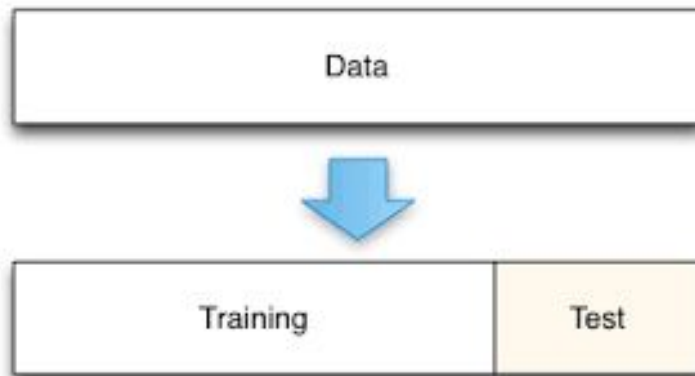


3. Logistic Regression

3.3 Technologies: System overview



4.1 Results: How to evaluate

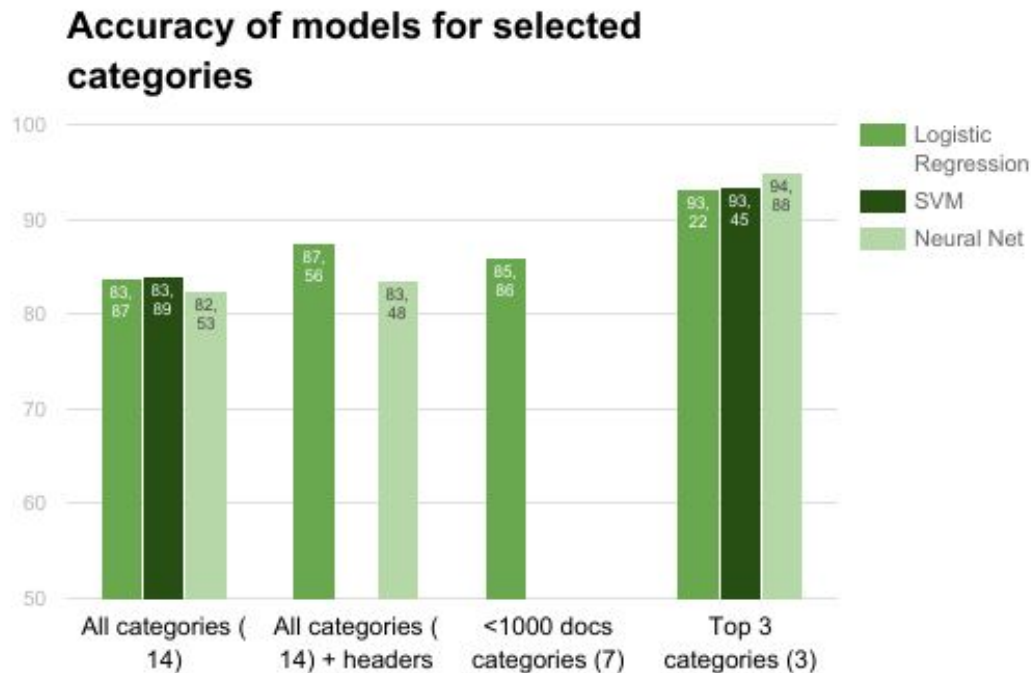


- First, give a system some **categorized documents** to **learn** on
- Then, ask the system **to categorize** some **unseen** documents whose true **category** we **know**

4.2 Results: Accuracy

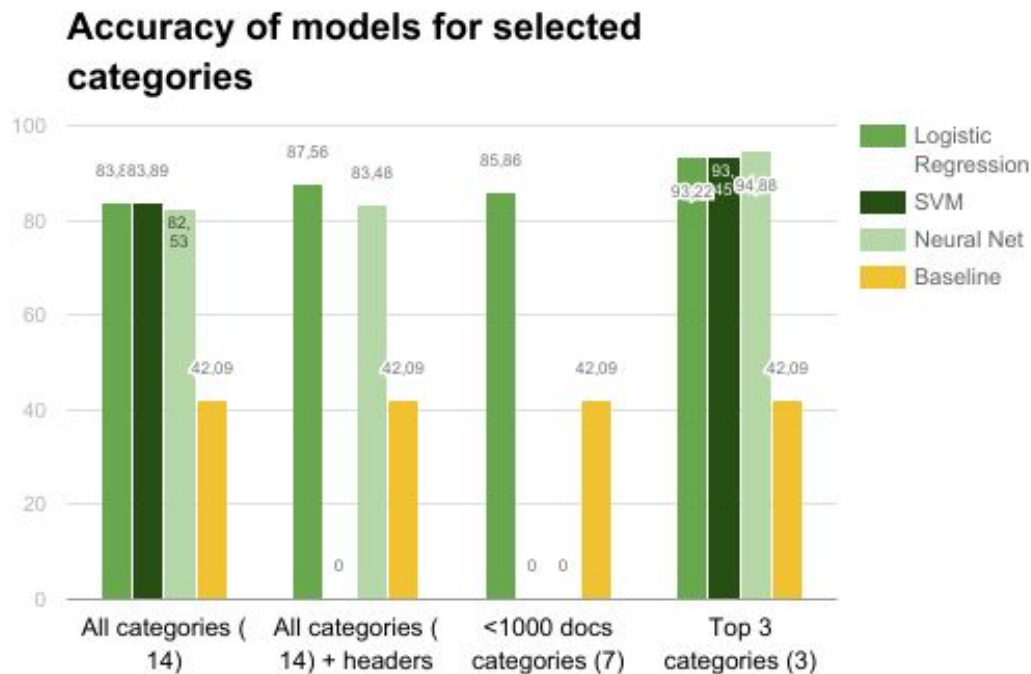
- **Percentage** of documents assigned to a **correct category**

$$\text{Accuracy} = \frac{|D_{\text{correct}}|}{|D_{\text{all}}|}$$



4.2 Results: Accuracy

- **Percentage** of documents assigned to a **correct category**



5:1 Use-cases: RHD books

- We've used our tool to score **books** from <https://developers.redhat.com/resources/#!type=book>
- **85** scored **books**
- content = abstract (- some are very short: 4 to 122 words, mean 27.9 words)

Keyword	Product	Publish Date
<input type="text" value="Filter by keyword"/>	<input data-bbox="705 540 937 556" type="text" value="Seled Product..."/>	<input type="text" value="Filter by Date"/>

85 results found

Sort results by and show results per page



Java Persistence with Hibernate

January 6, 2015

Persistence—the ability of data to outlive an instance of a program—is central to modern applications. Hibernate, the most popular provider of the Java Persistence standard, offers automatic and transparent object/relational mapping, making it a snap to work with SQL databases in Java applications. Java Persistence with Hibernate explores Hibernate by developing ..



Advanced JAX-WS Web Services


September 30, 2014

In this book you'll learn the concepts of SOAP based Web Services architecture and get practical advice on building and deploying Web Services in the enterprise. Starting from the basics and the best practices for setting up a development environment, this book enters into the inner details of the JAX-WS in...




5:1 Use-cases: RHD books

Apress®

 [Login](#) [Apress Access](#)

CATEGORIES SERVICES APRESS OPEN BLOG WRITE SHOP

Search 



© 2004

JBoss 3.2 Deployment and Administration

Authors: Kunnumpurath, Meeraj

[Download source code](#)

[Free Preview](#)

About this book

Deploying J2EE applications and configuring the application server are some of the most tricky, non-standard elements of J2EE development. But *JBoss 3.2 Deployment and Administration* will help you navigate configurative, administrative, and deployment tasks on the JBoss application server.

Meanwhile, JBoss is the leading open-source J2EE application server. It was voted the JavaWorld 2002 Editor's Choice for Best Application Server. Further, a TogetherSoft poll named JBoss the leading choice (at 43 percent) for development.

About the authors

Buy this book

▼ eBook

27,36 €

price for Czech Republic (gross)

Buy eBook

- ISBN 978-1-4302-0819-8
- Digitally watermarked, DRM-free
- Included format: PDF
- eBooks can be used on all reading devices
- Immediate eBook download after purchase

> Softcover

36,37 €

amq	0.00220724
bpmsuite	0.00152983
brms	0.0178493
cdk	0.000151682
datagrid	0.00215388
devstudio	0.00245392
eap	0.955179
fuse	0.00163482
mobileplatform	0.000561591
openshift	0.0000 553
rhel	0.00195283
webserver	0.0141839

5:1 Use-cases: RHD books

- We've used our tool to score **books** from <https://developers.redhat.com>
 - **85** scored **books**
 - content = abstract (- some are very short: 4 to 122 words, mean 27.9 words)
-
- **47** books (55.3%) scored **reasonably** towards all 14 categories
 - Inaccurate major category score: **19 books**
 - Inaccurate minor categories' score: **16 books**
 - Hardly decidable by expert observer: **6 books**

5:2 Use-cases: Non-exact match search

Keyword	Product
<input type="text" value="keyword"/>	<input type="text" value="JBoss Fuse"/>
Publish Date	
<input type="text" value="Filter by Date"/>	

1 results found

Sort results by and show results per page



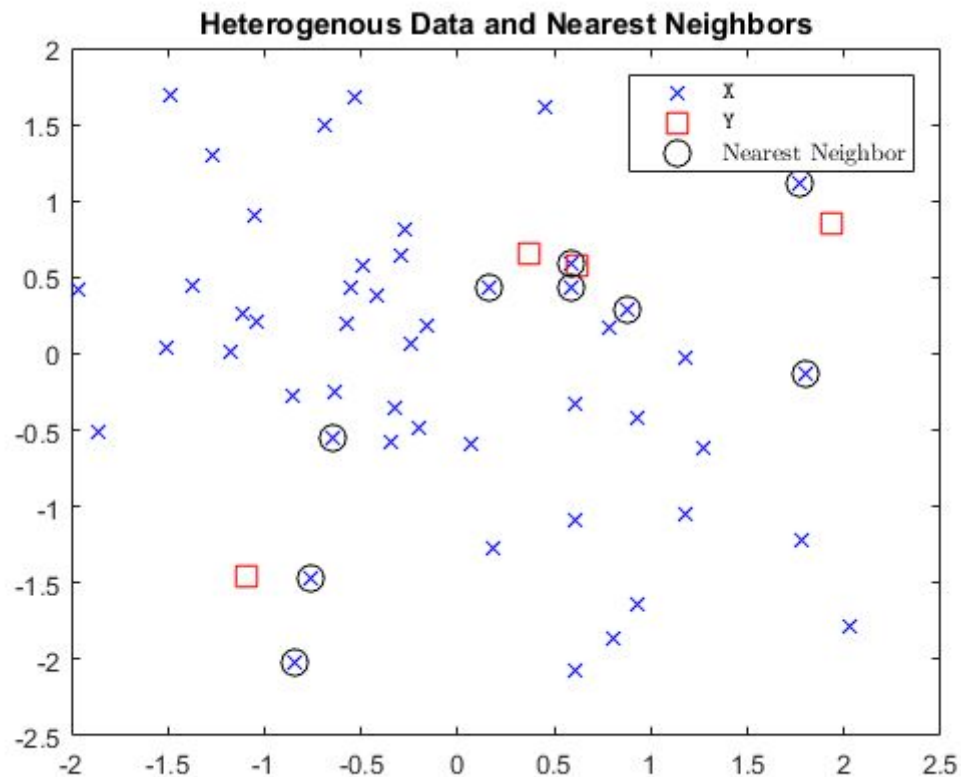
How can we block a container from starting

November 5, 2015

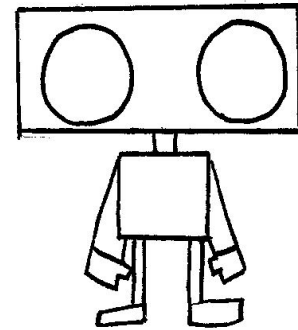
- We have three nodes fabric ensemble. We would like to have Prod DR containers in UAT enviroment, but we want to make sure that the DR containers don't accidentally start. We tried to modify the instances/instance.properties and prefix the "item.5.name = my_container_name" with block keyword for example "item.5.name = ...

Tags: fuse

5:3 Use-cases: Customer recommender



That all!



Michal Stefanik & Vlastimil Elias