



Design of Natural Language Processing Toolkits in the Industry: Lessons Learned and Takeaway Messages for Chatbot

Dec 6, 2021

Gias Uddin, Advanced Analytics, OSFI



OSFI
BSIF

The presentation is based on my work experience and opinions. It does not necessarily reflect the view of OSFI or my previous employers

Canada

I conduct research and build real-world impactful applications using data science tools and techniques

Academia



Queen's
UNIVERSITY

MSc, Computer Science
2006 - 2008

A mathematical trust model
for service-oriented software



McGill
UNIVERSITY

PhD, Computer Science
2010 - 2018

Automatically summarize sentiments and
reviews from social media data



UNIVERSITY OF
CALGARY

Assistant Professor (tenure-track)
Electrical and Software Engineering
2020 - Present

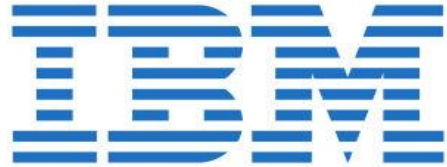
Founding director of Data Intensive Software Analytics
(DISA) lab to design responsible ML Software

Industry & Government



Data Engineer
2008 - 2010

Investment analytic
software development
using company stock
prices and balance sheets



Researcher & Software Engineer
IBM Watson Analytics
2011 - 2015

Automatically determine intents in the
textual queries of users and produce
output to answer to queries using NLP



Data Scientist
2016 - 2018

Processing of millions of
scanned mail labels from
Canada Post and generate
analytics (fraud detection)



BANK OF CANADA
BANQUE DU CANADA

Senior Data Scientist
Data and Statistics Office
2018 - 2020

Data governance and forecasting
model for TransUnion consumer
credit microdata

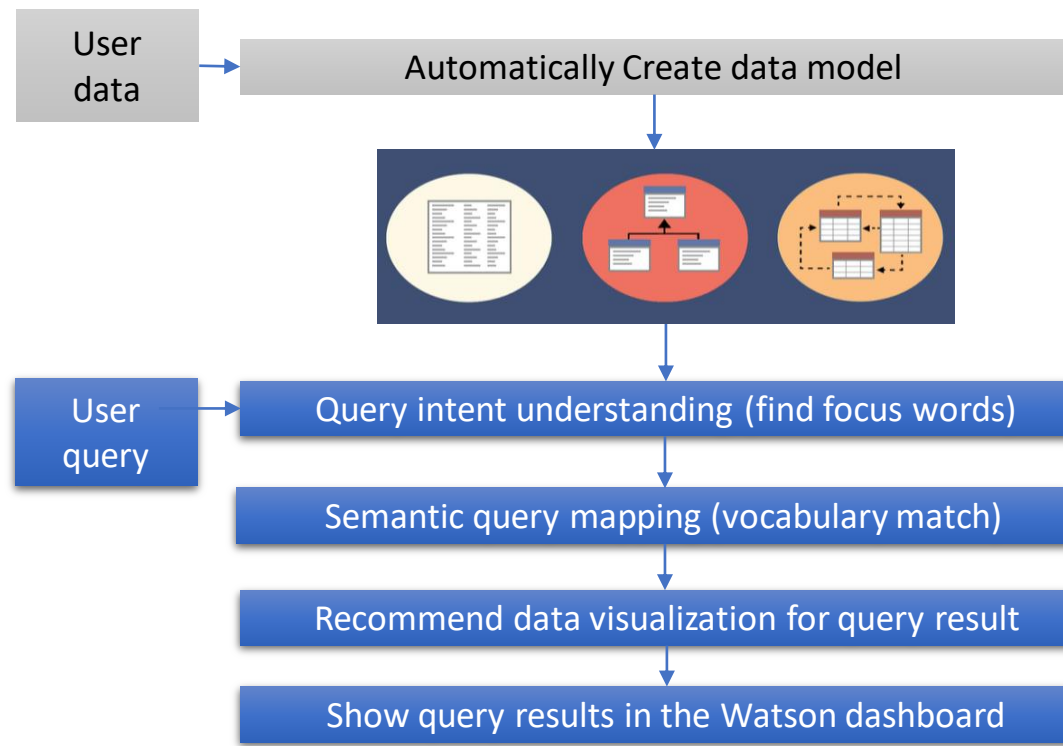


OSFI
BSIF

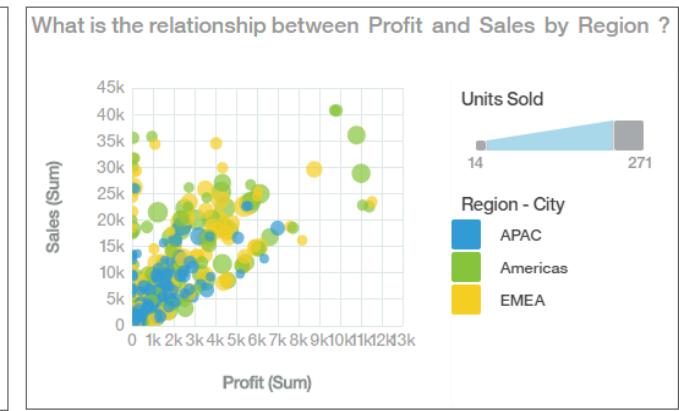
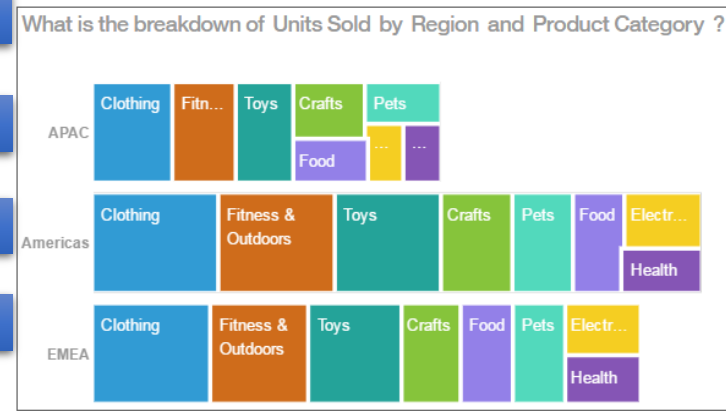
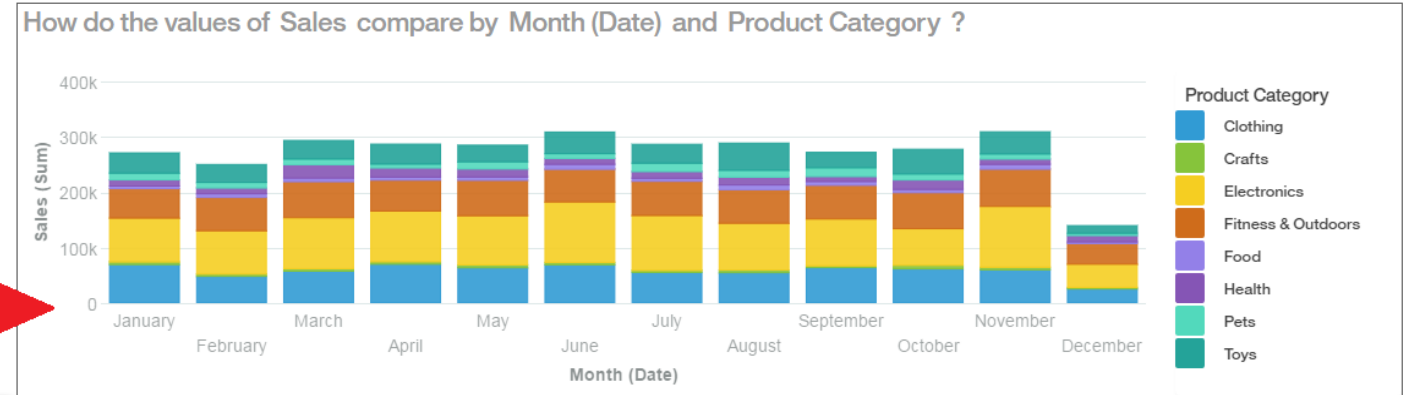
Manager
Advanced Analytics
2021 - Present

Design and development of
SupTech (Supervisory
Technology) tools using
AI/ML/NLP techniques

Automatic understanding and responding to user "Intents" in a natural language query



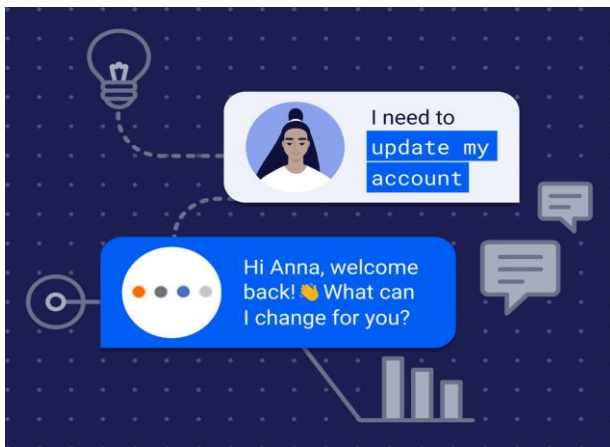
Product Sales Performance



Product development team consisted of hundreds of software developers, architects and product managers in several countries across two continents (North America & Europe)
 Required more than two years' of planning and development to build first consumer ready version

- Notes from IBM.** Watson Analytics was integrated into Cognos Analytics and would now be reflected in IBM Cognos Analytics Exploration hands-on demo.
- The IBM Accelerator site contains tutorials: <https://community.ibm.com/accelerators/?context=analytics&query=Get%20Started%20Tutorial>.
 - Cognos Analytics on the Cloud is now available on the Academic Initiative. For academic access, please follow the instructions at this site: <https://github.com/academic-initiative/documentation/tree/main/academic-initiative> and do the following steps:
 - Under General - How to register with the Academic Initiative.
 - Once #1 has been completed, They can go to the Topics drop-down, select Data Science, and then scroll down to access Software and then select Cognos Analytics in the cloud.

Lessons learned throughout the journey of developing various NLP tools and takeaway messages to build Chatbots



Tool

It is almost impossible to find a third-party tool that can do all the NLP tasks that can meet specific business objectives

It often requires extensive customization of the third party tools to tailor those to specific needs

Technique

ML model is only a part of many non-ML but NLP/data tasks in any textual analysis

Intent detection or any NLP task (e.g., classification) can be very much a domain-specific problem, requiring specialized ML/NLP models tailored to each domain

Operationalization

Operationalization of ML models often require a holistic change due to the integration of the ML models and the workflow/pipeline into the traditional software systems

Adoption of NLP tools needs to be human-centered by focusing on specific needs and by continuously updating the tools based on changing requirements

Responsible AI

There can be a fundamental change in approach between how a model is designed during exploration phase vs how it may be used in real-world use cases

Wrong data or wrong technique/tool in a Chatbot can have harmful consequences besides providing wrong outcome



Microsoft's chatbot Tay was shut down 16 hours after launch as it had learned slang and inappropriate language from Twitter trolls.

[https://en.wikipedia.org/wiki/Tay_\(bot\)](https://en.wikipedia.org/wiki/Tay_(bot))