Seton Hall University

eRepository @ Seton Hall

Petersheim Academic Exposition

4-2020

Data Science Meets Compliance

Christian Clarke

Follow this and additional works at: https://scholarship.shu.edu/petersheim-exposition

Part of the Artificial Intelligence and Robotics Commons, and the Data Science Commons

Recommended Citation

DATA SCIENCE MEETS COMPLIANCE

CHRISTIAN CLARKE

WHAT EXACTLY IS DATA SCIENCE?

- Massive influx of data & improved ability to properly handle it

- Blend of various tools, algorithms and machine learning principles with the goal to discover hidden patterns from raw data

- Primarily used to make decisions and predictions making use of predictive casual analytics, prescriptive analytics, and machine learning

  - Predictive Analytics: Using data to create a model that can accurately tell what will happen in the future (i.e predict if customers will make future payments)

  - Prescriptive Analytics: Using data to supply recommendations based on trends (i.e self-driving cars)

  - Machine Learning & Artificial Intelligence: The ability to do the above without the need for human interaction

WHAT EXACTLY IS COMPLIANCE?

- The process of making sure a company and employees follow all laws, regulations, standard, and ethical practices that apply to the organization

- The goal is to protect the business from any form of liability or litigation

- Should be integrated into all parts of the business

- Should be a seamless interaction with as few bumps and disruptions as possible

- Often dealing with an incredible amount of forms and applications

THE PROBLEM

- Large amounts of paperwork make it difficult to sift out compliance mistakes

- Delayed discovery of mistakes leads to delays in handling of mistakes

- Delays in handling of mistakes open companies up to litigation and liability

- Need a quick and reliable way to process paperwork and flag possible issues

  - Expense Reports filed from non-sanctioned locations

  - Payments made to non-sanctioned parties

  - Large expenditures

THE SOLUTION

- Incorporate geo-location data to identify types of locations

- Use Natural Language Processing tools to analyze report structure to identify possible inconsistencies

- Use Supervised Learning techniques to train models to automatically recognize anything that does not meet compliance standards

- Tune Machine Learning algorithm and incorporate it into cloud computing network to scan compliance data automatically