Seldon-summary

v1 mjd 24.apr.2022

SELDON

Automated/Synthetic intelligence generation from massive and diverse social media (SIG)

Applied/automated intelligent psychophysical analysis (AIPA)

It is possible to produce accurate information regarding the present and future psychological and physiological states of individuals and groups, including moods, temperaments, dispositions and intentions, through the analysis of digital media created and disseminated across electronic media including multiple types of social networks.

This information is produced by the analysis of features in three types of media: (1) specific media objects (e.g., photos, videos, non-photo images such as graphics and icons, audio messages and texts), (2) sequences of such types of media linked with the same individuals or groups and involving changes in time and location, and (3) comparisons among such media collected over sequences of time and in different locations, from among many different individuals.

This process of knowledge extraction makes use of multiple, parallel forms of machine learning including rule-based and statistically-based algorithms and systems. There are parallel, competing agents (methods) applied to the data streams and the results are integrated by the use of the same fundamental types of learning algorithms, thereby creating a set of probabilistic statements.

The resulting information derived from such massive media analytics is the basis for generating probabilistic inferences regarding the behaviors of both individuals and groups within different physical, psychological, social, and economic contexts. Such contexts are of two types: (1) actual historical situations, and (2) projected response to different types of stimuli and environmental conditions (including those of a socioeconomic nature).

This information is of value to the respective individuals and groups, for their own interest and use. It is also of value to client-subscribers (companies, institutions) wishing to make decisions based upon projected response and behaviors by their target audiences. This value extends through multiple markets and contexts including large-scale financial, political and other social-impact processes.

Objectives:

- Knowledge about individuals and groups
- Current-state and future-state (predicted, projected)
- Physical states (health-related)
- Psychological states
- Reasoning about behavior and receptivity to products and socioeconomic actions

Input types:

Digital media of all object types
images (photos, video, graphics)
sounds (voice messages, music)
texts (posts, blogs, documents)

Input sources:

Social networks
Linked/paired apps
Emails and other transmissions that are shared

Resources Employed (selected list)
Activity-based intelligence
Exploitation analysis and countermeasures
Cyber operations planning and CARVER methods
Alternative compensatory control measures
Asymmetric and symmetric encryption - conventional, novel and exploratory
---- including RSA, AES, Diffie-Hellman, El Gamal, PGP, Rijndael, Elliptic Curve (ECC, ECDSA)
Machine learning and evolution
---- Bayesian, Biometric, Genetic, Neural, Logic/rule-based, Quantum, Statistical, SVM, Thermodynamic
ETL, MPP, and VLDB data mining
Homomorphic encryption models and applications
Agent-based and botnet architectures and defenses
Anomaly and intrusion detection and countermeasures
Behavior modeling, profiling and predictive analysis
Insider threat modeling, detection and tracking
Industrial and financial espionage countermeasures
Vulnerability assessment and monitoring
Asset validation and protection
Risk factor investigation and evaluation

Also, very strongly employed:

VAE, GAN, Bayes, SVM, and Formal Logic Systems

Seldon incorporates a new integrative and bio-social approach toward building an architecture for machine intelligence that will have the foundational capabilities for innovation, invention, synthesis-thinking, discovery, and self-awareness that is capable of contributing particularly to prediction and forecast of behavioral dispositions and trends for very large populations (e.g., humans, but not only such) – in other words, a Prediction Engine that complements and extends the contemporary very large data collecting and aggregating capabilities of search engines and similar tools.

Thus, we have been engaged for years in research within what is commonly called artificial intelligence and in recent years “deep learning” and which we believe and assert is best termed

“Synthetic intelligence” (SI). The technical aspects of this involve the integration of multiple forms of machine learning, pattern recognition, data compression, theorem proving, and other aspects of intelligence that can be implemented through computers in a manner that creates a community of intelligent agents working in both asynchronous and synchronous parallel computations and cooperating with humans and other machines. This involves logic programming, neural networks, probabilistic reasoning, and technologies including specific forms of VAE, GAN, Bayes, SVM, and other paradigms.

Within our work, this synthetic intelligence research has been directed toward pattern detection and modeling that serves predictive analytics for application to extremely large and diverse data sets pertaining to population dispositions, trends and behaviors. This is known as the Seldon Prediction Engine (or simply, Seldon – a reference to a fictional character who embodied some of these ideas).

The Prediction Engine (“PE”) is for making predictions, forecasts, estimations, for the expected behaviors of people and other population-like entities on the basis of known and projected attributes of behavior and potential behaviors (e.g., beliefs). The PE aims to provide both simple and complex answers with respect to both abstract, statistical populations and also specific real-world, real-life people and social groups.

The POP is the populations – the people, the groups, both “real-world” and “avatars.” That is a universe of documents, like the URLs comprising a search engine. The ATT is the attributes, the characteristics, the actions, all types and qualities. The PE works by mapping ATT to POP and vice versa, and assembling internal models of how members (what make up the POP) interact according to their elements (what make up the ATT), and especially, what those members will tend and choose to do in the future, based upon expected and predictable changes in both the POP and the ATT.

All software components within the PE are designed to be also usable, wherever possible, within stand-alone applications that may include mobile apps, bots in messaging and control systems, and virtually anything else that can be usable for two primary purposes: (1) generate capital through revenues, and (2) generate additional information about the POP members and ATT elements, in order to automatically, in the background, constantly, seamlessly, and unstoppably, grow the knowledge within the PE. These applications encompass the set of IID (Intelligence information Domains) and within their operating space, the SIA (Sient Intelligent Applications). [1]

There is no singular formula, equation, or master algorithm for the data and knowledge acquisition and the learning that derives from such. Many competitive and collaborative systems are used and usable.

The Prediction Engine Technology (PET) within Seldon is the basis for a generalizable system of predicting future outcomes and trends from a variety of data sets. The PET architecture is based upon a *process model* of data acquisition, content analysis, feature extraction and knowledge generation. This model is formally based upon the concept of process algebras involving concurrent (parallel) dynamic (reconfigurable, reprogrammable) agents or actors that are implemented in software. These agent processes function in both cooperative and competitive relationships with one another in the

1 The SIA and one in particular, Debora, are described in another summary paper.

course of executing queries made in regard to future events that involve elements within accessible data sets.

Topology is very much at the heart of the Seldon logic. The foundation of this model is the concept of *dynamic spatiotemporal relations*. This concerns the changing relationships between measurable features of an object (e.g., a person's face or body, a machine, or something else for which an analysis is being made in order to predict its current state and some future probable states. According to this perspective, the interpretation of many features (such as within images or text expressions) is dependent upon preceding and subsequent variations within the observed feature ("temporal" dynamics), and also upon other concurrent (or temporally separated) features which may be either nearby or distant, relative to the observed feature ("spatial" dynamics).

What matters is not only the raw data of some image or text, but how it changes over time, and also what it is with at or near the time that it has been (created, uploaded/posted, etc.).

Many predictions of future events, such as population trends and outcomes, or individual behaviors and dispositions, can be influenced by these two major factors which may be considered to be "non-local" with regard to spatial and temporal observations that are generally associated with the subjects of the prediction queries:

More coming soon...

