

# Artificial Intelligence in Banking and Finance



ITM TwentyFirst University - 2017

# Life Insurance 101

**Tuesday, November 14th, 2017 at 2 PM ET**

1 hour CE for CFP, CTFA and FIRMA

*Michael Brohawn, CFP, CLU*



ITM TwentyFirst University - 2017

# TOLI Issues & Solutions

*2017 - The Year In Review*

**Tuesday, December 12th, 2017 at 2 PM ET**

1 hour CE for CFP, CTFA and FIRMA

*Michael Brohawn, CFP, CLU*



# ITM TwentyFirst University - 2018

- **The Life Insurance Purchase Process**
- **Life Insurance Policy Valuation**
- **Estate Planning Update for 2018**

**COMING SOON**  
**2018**



# Today's Presenters

**Cory Cates**

Chief Information Officer

**Cody Pobuda**

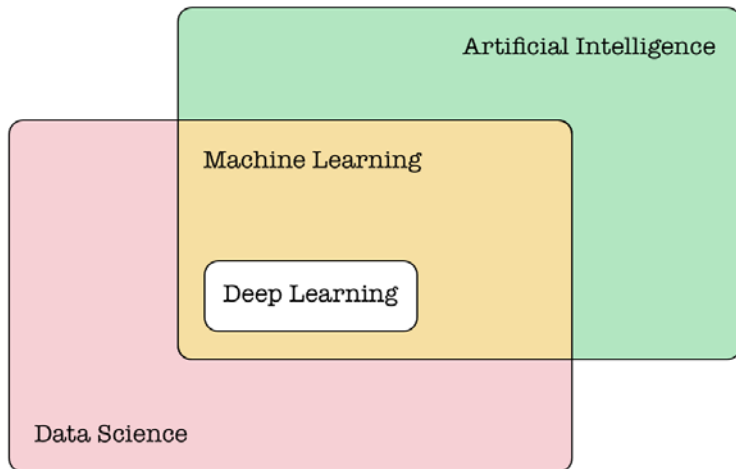
Manager of Technology



# What is Artificial Intelligence

Traditional computer programming techniques involve **defining logical rules** for different combinations of data.

Recent explosion in the field has been to use **training by example** rather than rules.



Shift from rules based models to machine learning and deep learning models allows handling of high uncertainty. Partial information, judgement calls, educated guesses.

As an example consider how children learn complex things - language, the recognition of objects and faces, or how to drive.

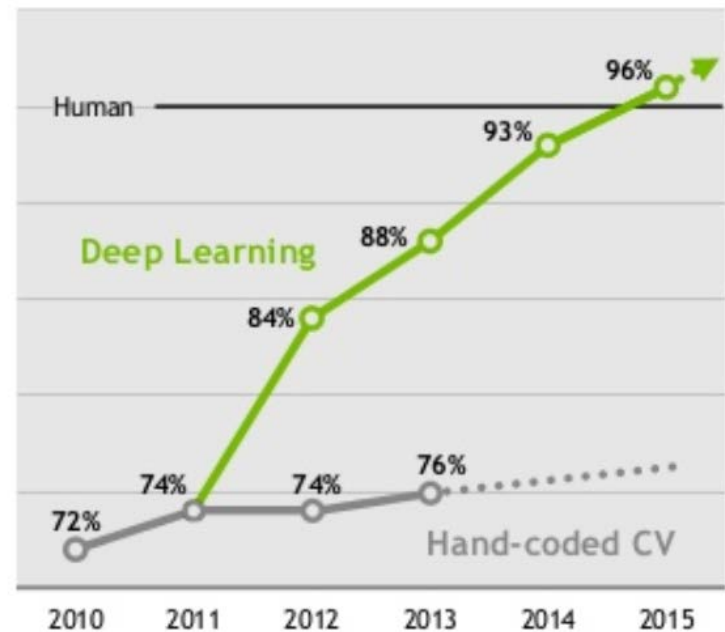
Difficult to determine the rules, but with lots of examples or experience it can be learned.

# The Power of Deep Learning

ImageNet Challenge – 1,000 categories, 1.2mm training photos, 100,000 test photos



ImageNet – Accuracy %

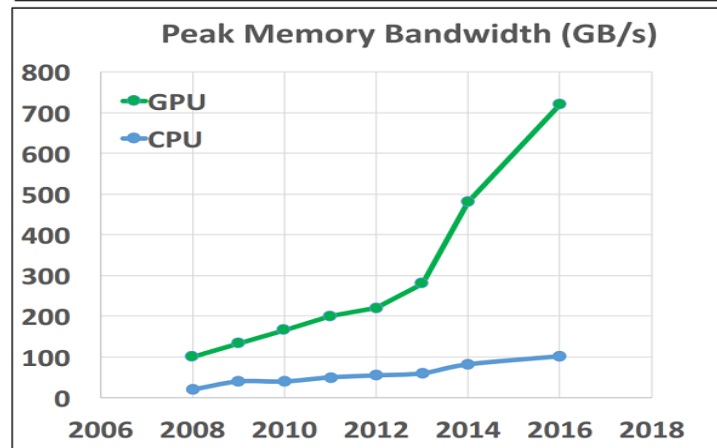
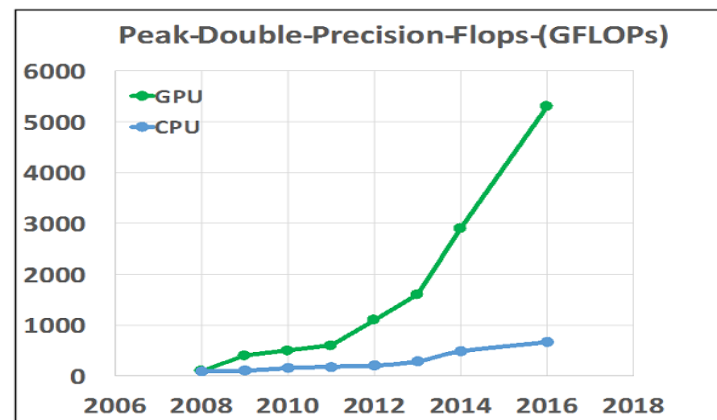


Deep Learning methods introduced in 2012 and dominate – by 2015 accuracy surpasses Human control group

# Why Now

Three key ingredients to recent explosive growth in machine learning:

- **Lots of examples**  
cheap data storage combined with larger data collection and more connected devices
- **Processing Power**  
Availability of GPUs to process larger amounts of data in parallel has opened up deep learning – slow training process now fast enough to get significant results
- **Strong Tools**  
Algorithms and advanced techniques and tools shared in the data science community – combination of scientific and open source mindset





# What Works Well

What we need:

- Good Questions – Needs to be an objectively answerable question
- Good Answers – Don't teach bad habits
- Lots of examples - Data is King

Watch out for:

- Overfitting – Computers can memorize the answers and will if you let them
- Lack of Transparency – Like human thought sometimes very difficult to understand how an answer was derived (How do you know that voice is Mike?)

## Examples

### Classification

Spam detection  
Credit Underwriting  
Customer Attrition  
Sentiment Analysis

### Clustering

Recommendations  
Market Segmentation

### Anomaly Detection

Credit Card Fraud  
Security – Breach Monitoring

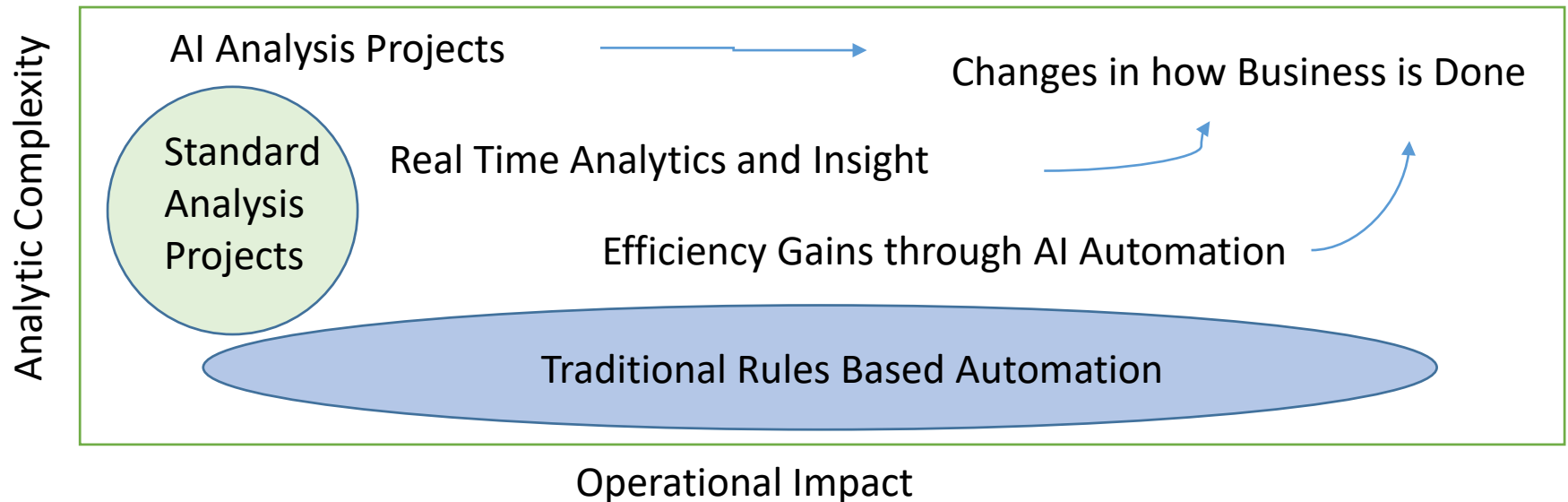
### Interface Improvement

Support – natural Language processing, chatbots  
Voice controls

### Real Time Decision Automation

Self driving cars

# Progression of Use



Common path for implementing AI tools is to find efficiency gains and extend or improve the capability of rules based automation. These tend to be improving a process that exists and is a quick win that pays for itself. Over time, especially when combined with higher level analysis to better understand what people are doing or how the market works – true transformation can happen that changes the way business is approached.

# Document Classification

- 10,000 to 12,000 Policy Documents in every month
- Our most important documents are:
  - Annual Statements
  - Quarterly Statements
  - Illustrations
  - Premium Due Notices
  - Grace Notices
  - Lapse Notices
- The above documents represent about 65-70% of all policy document volume.
- We would like to improve upon our current system which uses various rules and pattern matching to identify document types. This gives an accuracy around 50-60%, with the remainder going to humans to resolve.
- We would need to see a very high degree of accuracy on the above important document types.

# High-Level Approach – Starting Point

- OCR the documents
- Pass OCR text to a classification model
- That model will likely be a Deep Learning based.

# Challenges

- 38 document types, ~250 carriers, giving us as many as 10,000 variations
- Document noise
- Noise (misspellings) from OCR process
- Incorrect labels
- Composite documents (Change of Bene/Change of Owner)
- Many documents in single PDF (Policy Information/Premium Due Notice/Annual Statement)
- Ambiguous documents (could be two or more different types)
- Some documents you can't even get the humans to agree
- Uneven Distribution of documents, far more Premium Due Notices than Grace or Lapse Notices

# Solution

- OCR the documents
- Pass OCR'd text to many models
- Models are weighted more heavily towards Grace and Lapse if model type allows, tried to strike a balance
- Results from many models are then ensembled and feed to a final classification model
- Results from those models are also feed to a different model to determine if the document should be reviewed
- Hand reviewed many documents to check why the model was getting it wrong.

# Results

- Overall:
  - Correct: 94.94%
  - Review-Correct: 3.45%
  - Review-Incorrect: 1.02%
  - Incorrect: 0.59%
  
- Important Documents:
  - Correct: 97.40%
  - Review-Correct: 1.96%
  - Review-Incorrect: 0.37%
  - Incorrect: 0.27%

# Key Takeaways

- Spent too much time in hyperparameter optimization within one model type (Deep Learning) and should have started trying other model types earlier on.
- Deep Learning is not a panacea. It's showed a lot of progress in the computer vision space, but is not necessarily going to be the best approach for everything.
- Many good models ensembled together will usually outperform one really good model. (Wisdom of Crowds)
- Documents can be many things at one time, so your applications and models should be designed to accommodate that if at all possible.



# Future Work at ITM TwentyFirst

## **Continue Efficiency Gains**

Data Collection, Categorization and Extraction

Increase Accuracy, Improve Speed, Improve Ability to Scale

May allow us to do work and collect data that is not currently feasible

## **Predictive Analytics**

Policies going into Grace

Carriers raising COI

Future Health conditions or changes in LE

## **Better Identification and Matching**

Obituary Matching

Predicting and Finding when people move (Locate Service)

More granular underwriting for LE

# Future Trends in Finance

Support Interface – chatbots and automated voice

Security – Even more Fraud Detection – extend to Login and other Security – could extend to biometric matching

Fintech trend – making credit risk decisions using more information – insurance underwriting is following same trend

Could be more trade off of giving more information for improved underwriting – similar to chips in cars – wearables for health or life insurance, spending habits for credit risk – more personalized UW

Personalized Product suggestions and advice

More analytics / scenarios – more information to deliver complex analytics and uncertainty to simpler actions for consumers

# Questions?

