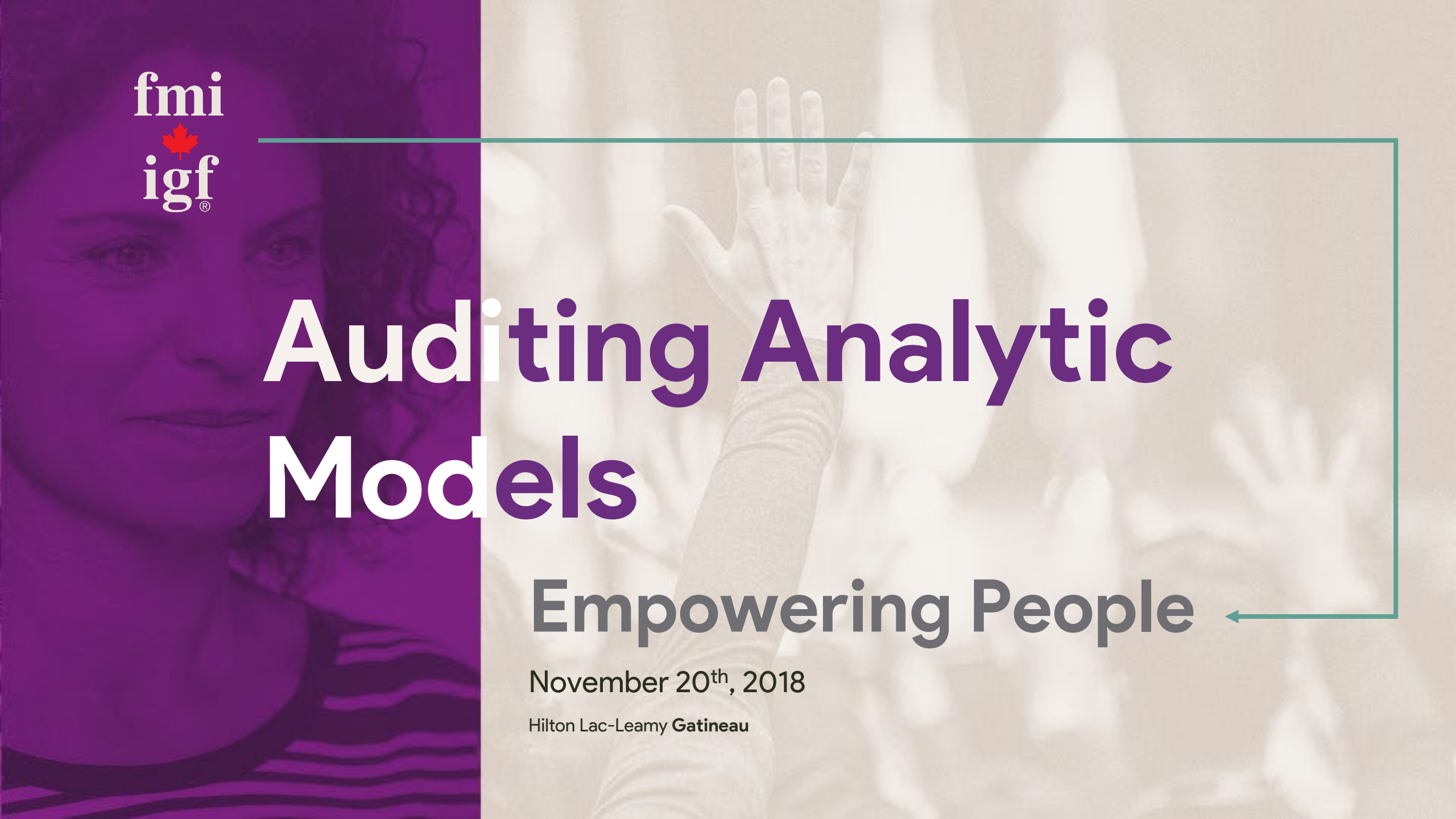


Auditing Analytic Models

Empowering People

November 20th, 2018

Hilton Lac-Leamy **Gatineau**

Why now....

"For internal audit to continue to deliver value, it needs to understand the impact of these technology innovations on their organizations. Boards and Executives expect internal audit functions to have a point of view on new technologies and to be able to provide advice on how the organization should embrace them while managing the risks and implementing strong controls."

- PWC 2018 State of the Internal Audit Profession Study

"How can CAEs upskill the internal audit activity to be ready for the challenge? The first step is recognizing that new skillsets are required. Collectively, the internal audit activity must have a sufficient understanding of AI, how the organization is using it, and the risks that AI represents to the organization. The CAE must be able to communicate this understanding to senior management, the board, and the audit committee. A good place to start is with The IIA's thought leadership on AI, and The IIA's supplemental guidance on topics like big data and talent management."

Outline

- *Quick Introduction to Advanced Analytics*
  - *What is Data Analytics*
  - *Basic Modelling Techniques*
    - *Regression analysis*
    - *Classification*
  - *Advanced Analytics*
    - *Supervised vs Unsupervised*
    - *The Neuron – the foundation of AI*
    - *Neural Networks (3 main types you should know)*

- **Auditing Analytic Models**
  - Introduction
  - Analytic Model Categories
  - Analytic Controls
  - Risk Assessment
  - Audit Approaches
  - Controls and Tests
  - Competency Profile and Training Plan

A Quick Introduction to Advanced Analytics

Before that…A quick introduction to me…

Allan Sammy,

Director, Data Science and Audit Analytics, Canada Post

- M.Sc. Predictive Analytics (Northwestern)
- CPA, CIA
- Director, Fraud Risk Management, Ontario Lottery and Gaming
- Director, Internal Audit, Canadian Air Transport Security Authority
- Forensic Accountant, Deloitte
- Commercial Crime Investigator, Royal Canadian Mounted Police (RCMP)

How I got into Analytics…….

What is Data Analytics?

- Data analytics is the pursuit of extracting meaning from raw data using specialized computer systems. These systems transform, organize, and model the data to draw conclusions and identify patterns.
- Benefits:
  - Cost Reduction
  - Faster Better Decision making
  - New Products and Services

These big trends are not that hard to spot (they get talked and written about a lot), but they can be strangely hard for large organizations to embrace. We're in the middle of an obvious one right now: machine learning and artificial intelligence.

- Jeff Bezos, 2017 Letter to Shareholders

What is a Model?

An analytical model is simply a mathematical equation that describes relationships among variables in a historical data set. The equation either estimates or classifies data values.

- Basic Modelling (Need to understand on a functional level)
- Advanced Analytics (Need to understand on a conceptual level)

Computers can only do math….

Basic Modelling Techniques

The primary basic modelling techniques that we are interested in are:

Linear Regression

- Regression analysis is used to understand which among the independent variables are related to the dependent variable, and to explore the forms of these relationships.
- Uses:
  - Model expected sales, expenses for outlier detection

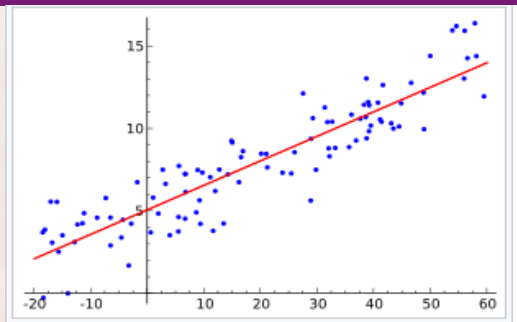
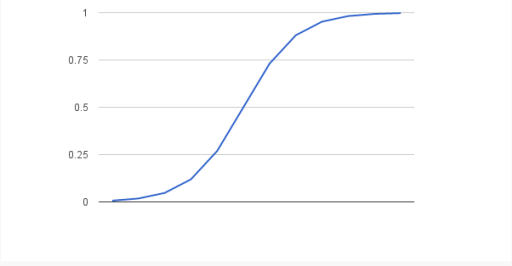


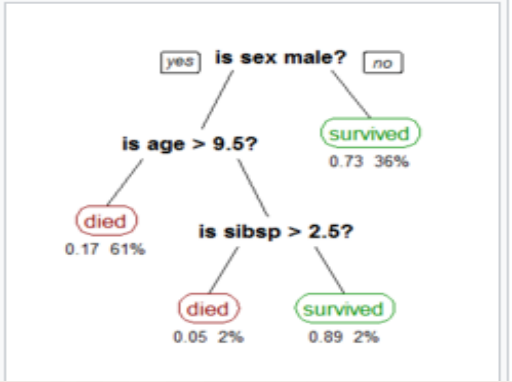
Illustration of linear regression on a data set.



Logistic Function

Classification

- Classification is the problem of identifying to which of a set of categories (sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known.
  - Logistic regression
  - Decision trees



Advanced Analytics

Important concept : Supervised vs Unsupervised Learning

- Supervised : The model is trained on “labelled: data

✓ Known fraudulent transactions

✓ Pictures of dogs and cats

- Unsupervised: The model is trained on “unlabelled” data

✓ Transactions, some of which may be fraudulent

✓ Bunch of pictures of animals, people, things

Advanced Analytics

The Neuron – The Foundation of AI

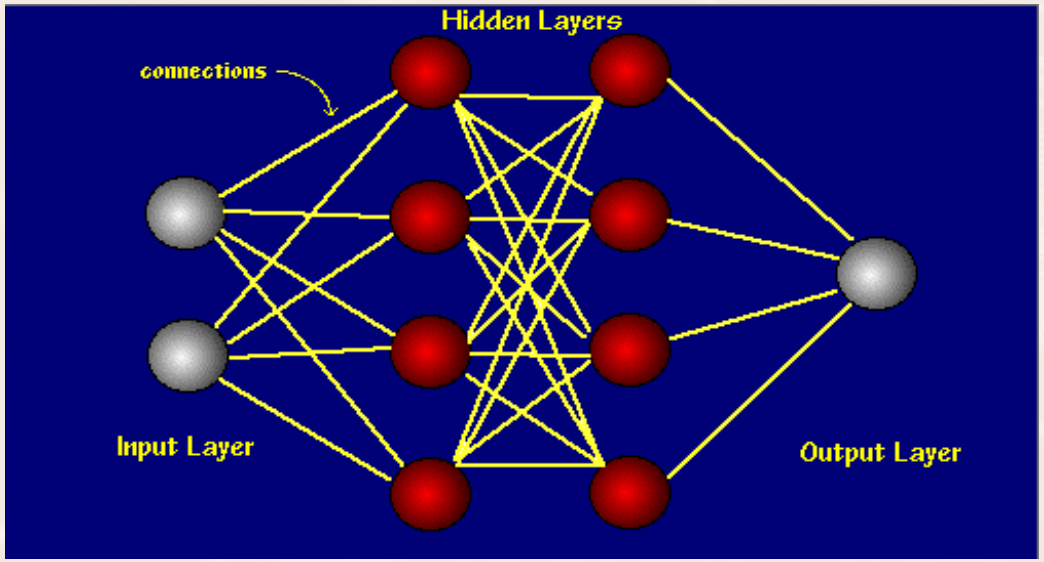
- Based on the brain
- Takes a series of inputs, assigns a weight and decides if it should “fire” or not.
- Groups of neurons work together in a network
- We are all the subject of various NN’s everyday….

Advanced Analytics

- Neural Networks

- Neural networks are typically organized in layers. Layers are made up of a number of interconnected 'nodes' which contain an 'activation function'. Patterns are presented to the network via the 'input layer', which communicates to one or more 'hidden layers' where the actual processing is done via a system of weighted 'connections'.

- What does that mean???



Advanced Analytics

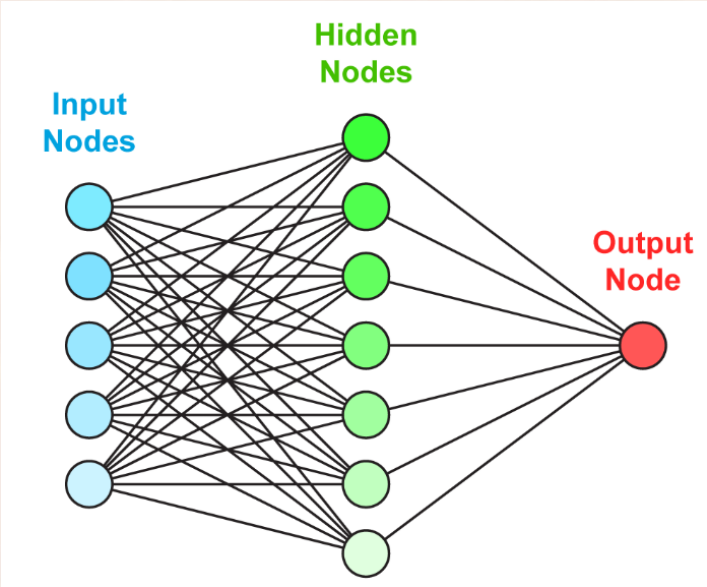
Kicking a soccer ball…..

1. Take a good touch
2. Look Up (optional).
3. Look at the Ball
4. Place your foot.
5. Swing your arms.
6. Bring back your kicking leg.
7. Lock your ankle
8. Keep your body straight.
9. Angle your body
10. Kick the ball with the correct part of your foot
11. Follow through
12. Follow up (optional)

Advanced Analytics

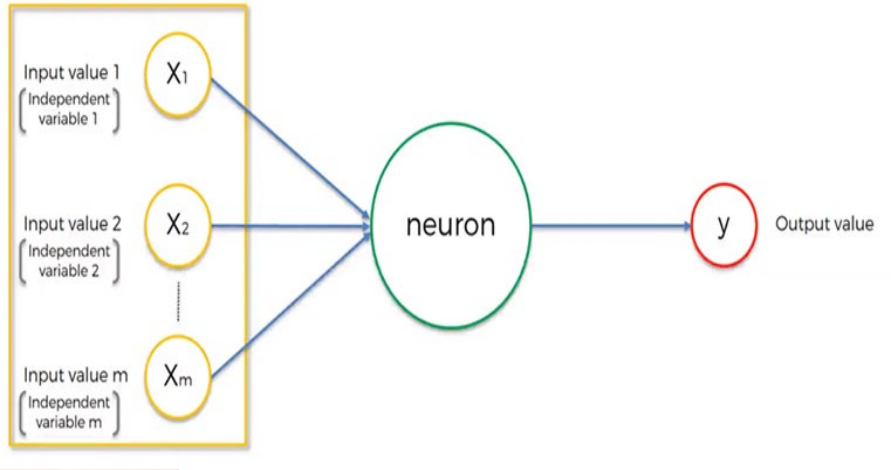
- How do you learn to kick a soccer ball with all those variables involved?

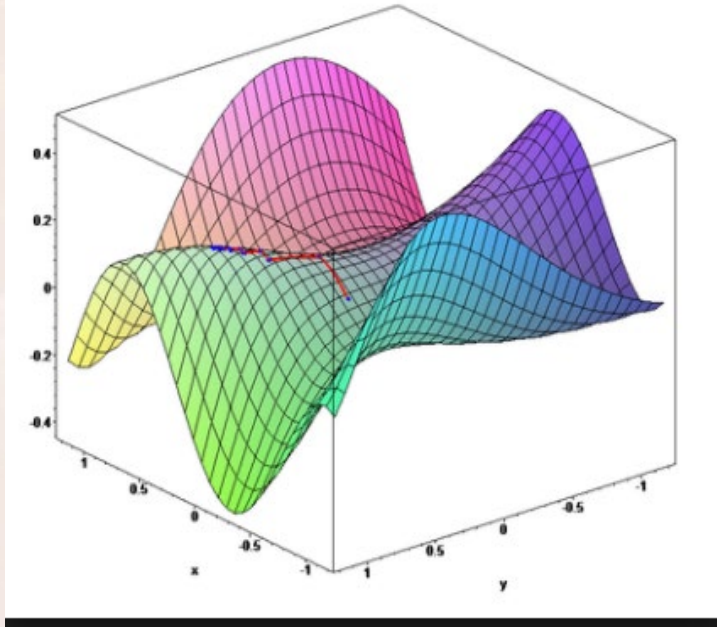
- Start somewhere (foot placement, arm swing, body angle, etc.) and take a shot.
- Look at the result, if not satisfactory, adjust a couple of the variables and try again.
- Repeat many, many times until you have a consistent accurate result.
- The combination and weighting of variables that works for you is your “model” of how to kick a soccer ball.
- Neural Networks work the same way….



Advanced Analytics

- A neural network starts with arbitrary weights and runs thousands of examples through the network.
- The network then checks the score (how many it got right).
- Adjusts the weights through a process called backpropagation (chain rule calculus)
- Runs thousands of examples and checks the score again (*repeat this process hundreds/thousands of times*)
- Every result is plotted on a loss surface (multidimensional plane)
- Each time the weights are adjusted, the network checks to see if it is moving in the right direction – towards the lowest point. This process is called *gradient descent*
- This kind of NN we have been discussing is called a feed forward NN or multi layer perceptron.

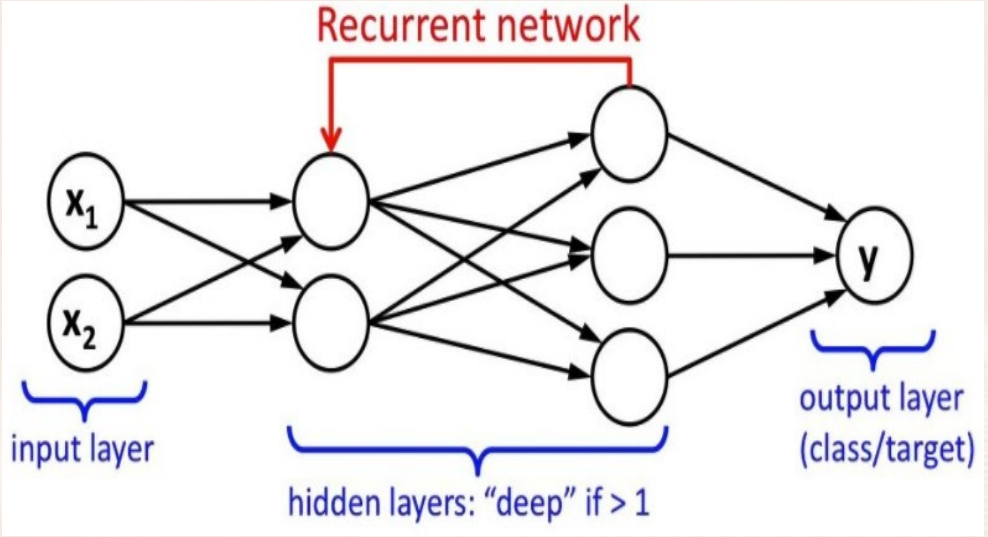




Advanced Analytics

Recurrent Neural Networks:

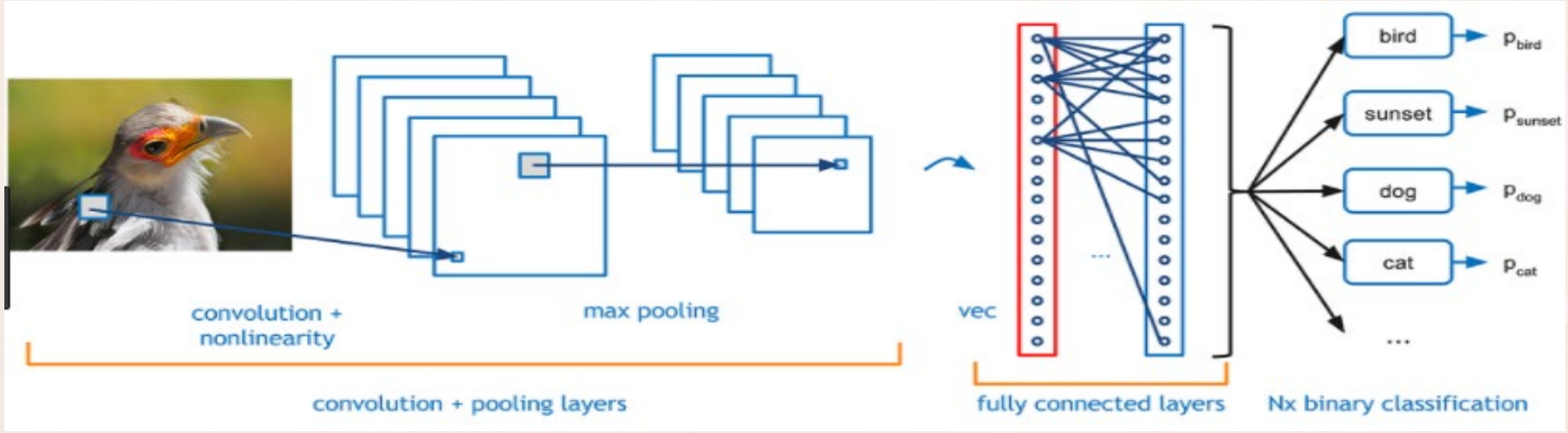
- Similar to feed forward network, however; some results from subsequent layer are fed back to previous layer and the weights are updated accordingly
- Used for forecasting/time series analysis, text analytics/prediction, translation
- The future state can be fed back to the present to refine the model. Ex guess the next word:
- “We are going to run a…” race? simulation? competition? program?



Advanced Analytics

Convolutional Neural Networks:

- Used primarily for image recognition/computer vision
- System passes several filters across an image and records the weighted value of each frame
- Combines these filters into a smaller picture and runs the process again
- Flattens the matrix into a series of inputs , then behaves similar to a FF NN

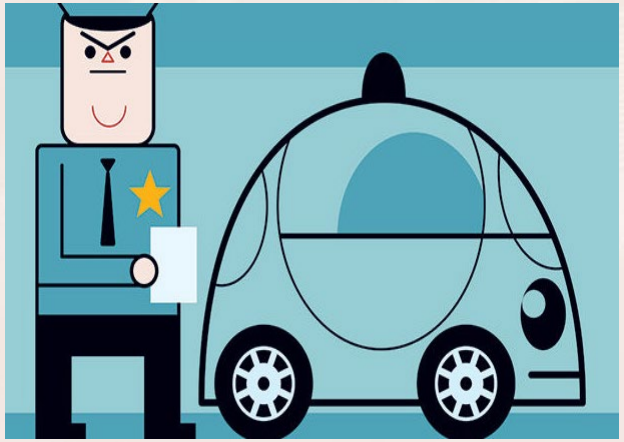


Auditing Analytic Models

Auditing Analytic Models – Introduction

- Organizations around the world are spending considerable money to build or buy analytic models and analytics capability to take advantage of big data, machine learning, and artificial intelligence (AI) technologies.
- These models have made their way into every aspect of business and are being relied on as decision support— and, in the case of machine learning and AI, actually making the decisions — for issues such as:

- Determining the probability of default for potential borrowers (corporate and individual).
- Evaluating new employees' probability of success and tenure with the organization (from professional athletes to salespeople).
- Forecasting success and return on investment for new marketing initiatives.
- Making product mix and store location decisions.
- Making life-and-death decisions in self-driving vehicles….



Auditing Analytic Models – Analytic Model Categories

There are three main categories of analytic models: Descriptive, Predictive, and Prescriptive.

Each category can provide an organization value and strategic insight

Descriptive

These models allow organizations to condense big data into smaller, more digestible pieces of information. Typically, organizations that use analytics meaningfully have mountains of raw data at their disposal. Descriptive analytics enables an organization to summarize that data and determine what really happened. Most analytics in use are descriptive: sales breakdowns, social media likes and followers, ratings, and reviews.

Auditing Analytic Models – Analytic Model Categories

Predictive

The next level up in data analysis, predictive analytics uses a variety of statistical, modeling, data mining, and machine learning techniques to study recent and historical data, enabling analysts to identify patterns and correlations in the data. Based on these identified patterns and correlations, analysts can create a model of the future results given selected inputs. For example, based on certain borrower characteristics, a bank may use a predictive model to forecast its amount of loan defaults.

Prescriptive

The highest level of analytics, prescriptive analytics recommends one or more courses of action and shows the likely outcome of each decision. Unlike a predictive model, a prescriptive model shows multiple future scenarios based on a decision the organization makes today. Prescriptive analytics requires a predictive model with two additional components: actionable data and a feedback system that tracks the outcome produced by the action taken. An example of prescriptive analytics would be a casino floor product mix optimization model that predicts revenue gains given various game configurations.

Auditing Analytic Models – Analytics Controls

- Today's organizations have billions of dollars riding on the accuracy and performance integrity of analytic models. With model performance becoming a strategic enabler, organizations need to manage the risks associated with analytics.
- Internal auditors need to assess the analytics controls for the models their organizations rely on to support decision-making.
- To effectively manage these risks and move beyond simple financial model or spreadsheet auditing, organizations need a system of controls around analytic model development.
- These analytics controls provide checks and balances around model selection, validation, implementation, and maintenance.

Auditing Analytic Models – Analytics Controls

The objective of analytics controls is to ensure that:

- Analytics personnel have the appropriate skills and training.
- Input data is appropriate, complete, authorized, and correct.
- Model selection procedures are documented and justified.
- Model validation and testing have been conducted in accordance with scientific principles.
- Outputs are accurate, complete, and being used by the business as intended.
- The model is refreshed and re-evaluated periodically.
- The organization maintains a record to track the processing of data from input, to processing, to the eventual output.

Auditing Analytic Models – Analytics Controls

There are several types of analytics controls:

Skills controls provide assurance that data analytics personnel are competent and sufficiently trained in relevant analytics methods.

Business use controls provide assurance that the model addresses the intended business objective.

Data controls are used mainly to check the integrity of data entered into an analytic model.

Model selection controls ensure model selection is appropriate and reasonable to provide decision support.

Auditing Analytic Models – Analytics Controls

Model validation controls address what is done to ensure the model output is reasonable and accurately reflects the underlying nature of the input data.

Output controls provide assurance that the model output is presented and used in an appropriate and justified manner to ensure it remains consistent and correct.

Maintenance controls address the need to re-evaluate and refresh analytic models periodically to ensure they are still relevant in the current environment.

Auditing Analytic Models – Risk Assessment

RISK ASSESSMENT

To add value to organization wide analytics control risk assessment activities, internal auditors should define the universe of analytic models and supporting technology (modeling software, data services, etc.). They also should summarize the risk and controls using the risk and control matrices documented during the risk assessment process.

Auditing Analytic Models – Risk Assessment

Next, internal auditors should define the risk factors associated with each analytic model by answering questions such as:

- Does the model support a regulatory requirement?
- How complex is the model type?
- How effective is the design of analytics controls?
- Is the model prepackaged (off the shelf) and customized or developed in house?
- Does the model support more than one critical business process?
- How is the data processed by the model classified (e.g., public, private, confidential, etc.)?
- How frequently are changes made to the model?
- How complex are those changes?
- What is the model's financial impact?
- How effective are the IT general controls residing within the application (e.g., change management, logical security, and operational controls)?

Auditing Analytic Models – Risk Assessment

Once they have answered these questions, internal auditors should weigh all risk factors to determine which risks need to be weighed more heavily than others . From there, they should determine the right scale for ranking each application control risk by considering qualitative and quantitative scales, such as:

- ❑ Low, medium, or high control risk.
- ❑ Numeric scales based on qualitative information (e.g., 1=low impact risk, 5=high-impact risk; 1=strong control, 5=inadequate control).
- ❑ Numeric scales based on quantitative information (e.g., 1=less than $50,000 and 5=more than $1 million).

Auditing Analytic Models – Risk Assessment

ASSESSING MODEL RISK

The example below of an analytic model control risk assessment uses a qualitative ranking scale (1=low impact or risk and 5=high impact or risk). Composite scores for each model are calculated by multiplying each risk factor and its weight in the model and adding the totals. For example, the composite score of 375 on the first line is computed by multiplying the risk factor rating times the specific model rating [(20 x 5) + (10 x 1) + (10 x 5) +...].

For this example, the internal auditor of a gaming company may determine that the analytic model control review will include all models with a score of 200 or greater.

ANALYTIC MODEL RISK FACTOR WEIGHTING

	20	10	10		10	10	10	15	15	
Model	Model Supports Regulatory Requirement	Degree of Model Complexity	Prepackaged or Developed		Model Supports More Than One Critical Business Process	Type of Data Processed by the Model	When Model Was Last Refreshed	Financial Impact	Effectiveness of the IT General Controls	Composite Score
Direct mail	5	1	5		5	3	3	5	2	375
Game mix	1	1	2		1	1	1	4	2	170
Lottery win	5	2	2		1	5	1	1	1	240
Anti-money laundering	5	3	5		1	5	5	5	2	395
Workforce	5	1	1		1	1	1	3	2	225

Auditing Analytic Models – Audit Approaches

If the analytics team uses a recognized methodology for model development such as the Cross-Industry Standard Process for Data Mining (CRISP-DM) or some other widely accepted system, then internal auditors should consider auditing to that standard. In addition, some organizations have established a model risk management function. Internal audit can audit that area using similar methodology to that applied to other compliance functions.

For organizations whose analytics teams do not use a prescribed model development methodology, there are two approaches auditors can use to audit analytics controls: the Integrated Model Review Methodology (IMRM) and the Stand Alone Model Methodology (SAMM). These methods apply CRISP-DM principles in an internal audit context.

Auditing Analytic Models – Audit Approaches

IMRM: This approach can be used to evaluate model risk by examining all the business processes that feed or are dependent on the model being reviewed. When using the IMRM, internal auditors should include within the review's scope all the organization's systems that are involved in the model under review and whether the implementation of the model is consistent with the organization's analytics strategy. In other words, the auditor needs to include within the review's scope the separate processes that make up the different components of the model cycle.

The auditor then can identify the inbound and outbound interfaces within the model and complete the scoping activity. For example, when auditors review a marketing campaign response model, they would scope in survey methodology and data collection processes, customer segmentation processes (inputs), and marketing decisions made based on model output.

Auditing Analytic Models – Audit Approaches

Using the IMRM approach typically devotes more audit resources to those analytic models that affect a larger portion of the organization's operations. To use the IMRM effectively, auditors need to understand the business processes surrounding the use of the model being reviewed and how data flows into and out of the model.

SAMM: The alternative approach, the SAMM, is used when the auditor wants to review the controls within a single model. The SAMM is useful for new models or when audit resources are limited. Essentially, the auditor is verifying that the model, itself, has appropriate controls and performs the intended function. It does not provide assurance as to whether the organization is using the model output effectively or whether the model inputs are valid

Auditing Analytic Models – Audit Approaches

Although SAMM is effectively a subset of the IMRM, internal auditors should clearly specify which methodology they are applying so that management and the audit committee know the extent to which they can rely on the results.

Auditing Analytic Models

IT'S STILL INTERNAL AUDITING

Although many auditors may be unfamiliar with analytic models, machine learning, and AI, the fundamentals of internal auditing remain the same. As with all new technologies and processes that organizations have embraced, internal auditors have a responsibility to learn how analytic models can be useful in their work and adapt their methods to serve their stakeholders.

Auditing Analytic Models

Controls and Tests

Auditing Analytic Models – Controls and Tests

Within the seven analytic control domains — skills, business use, data, model selection, model validation, output, and maintenance controls – there are specific controls that internal auditors can test. Auditors can use the Integrated Model Review Methodology to test all the controls listed in the tables below. With the Stand Alone Model Methodology, only data controls, model selection controls, and model validation controls are tested.

Auditing Analytic Models – Controls and Tests

Skills Controls (AC 1) — Provide assurance that data analytics personnel are sufficiently trained in relevant analytics methods.

Domain	Control	Possible Tests
AC 1.1 Analytic Competency	Analytics staff have appropriate baseline educational qualifications and/or sufficient experience in data science and data analytics for their role and level.	▪ Review employee files and assess level of analytic competency. ▪ Compare education/ experience at each level with established corporate competency levels, if available.
AC 1.2 Skills Maintenance	The analytics department has sufficient training budget and ensures employees' skills are kept up to date.	▪ Calculate training spend per employee per year and compare with established norms/benchmarks.
AC 1.3 Skills Deployment	Employees with the appropriate skills and knowledge are assigned to projects that suit their competencies.	▪ Cross reference results of AC 1.1 with type of analytic model under consideration and assess appropriateness of fit (e.g., a junior analyst should not be the sole resource assigned to a high value/high visibility model).

Auditing Analytic Models – Controls and Tests

Business Use Controls (AC 2) — Provide assurance that the model addresses the intended business objective.

Domain	Control	Possible Tests
AC 2.1 Business Objectives	Analytics personnel understand business requirements and have documented success criteria.	▪ Review documentation detailing the analytic model success criteria and ensure that business owners and analytics personnel have signed off on the criteria. ▪ If that documentation is unavailable, interview business owners and analytics personnel separately and assess the level of agreement between business owners' expectations and analytics personnel's understanding of the business problem and success criteria.
AC 2.2 Economic Justification	Use of analytic models is an appropriate and cost-effective method to address the business problem.	▪ Obtain cost-benefit analysis from business owners. ▪ If document is unavailable, conduct cost-benefit analysis.
AC 2.3 Model Applicability	The type of model selected is appropriate for the business situation and objective.	▪ Review the list of analytic model and assess choice of model in relation to the business problem. ▪ If model choice seems inappropriate, obtain rationale from analytics personnel.

Auditing Analytic Models – Controls and Tests

Data Controls (AC 3) — Provide assurance on the integrity of data entered into an analytic model.

Domain	Control	Possible Tests
AC 3.1 Data Acquisition	Data sources are reliable and appropriate for the model under review.	▪ Review data source and assess level of quality control and reliability (i.e., open source data vs. purchased data). ▪ Verify accuracy of data by corroborating a sample with other sources, if available (e.g., economic indicator data from a private data provider can be validated against Bank of Canada website data). ▪ Interview analytics personnel to determine reasons that the data set was selected for this model.
AC 3.2 Data Manipulation and Completeness	Model input data has been cleaned and post-verified. Input data transformations are documented and post-verified. Merged/aggregated data has been post-verified.	▪ Obtain documentation relating to all procedures used to clean raw data. Obtain evidence of post verification (i.e., key summary statistics are the same pre and post cleaning). ▪ Obtain documentation relating to any variable transformation procedures. Assess rationale and reasonability of data transformations. ▪ Obtain documentation related to data joins/links and verify procedures used to ensure data integrity.

Auditing Analytic Models – Controls and Tests

Model Selection Controls (AC 4) — Provide a means to ensure model selection is appropriate and reasonable to provide decision support.

Domain	Control	Possible Tests
AC 4.1 Variable Selection	Rationale for variable inclusion in the model is justified and documented.	▪ Obtain documentation regarding variable selection process. ▪ If documentation is not available, interview analytics personnel regarding variable selection process and rationale. ▪ Confirm with business owners that variables selected for inclusion in the model are appropriate.
AC 4.2 Model Choice	The choice of model is appropriate to the business problem and justified.	▪ Obtain documentation regarding model selection process. ▪ If documentation is not available, interview analytics personnel regarding model selection process and rationale.
AC 4.3 Model Assumptions	Assumptions around model selection (i.e., market conditions, proxy variables) are reasonable and supported by evidence and/or business owners.	▪ Obtain documentation regarding model assumptions. ▪ If documentation is not available, interview analytics personnel regarding model assumptions and rationale. ▪ Verify that business owners understand and agree with analytics personnel's assumptions.

Auditing Analytic Models – Controls and Tests

Model Validation Controls (AC 5) — Address what is done to ensure that the model output is reasonable and accurately reflects the underlying nature of the input data.

Domain	Control	Possible Tests
AC 5.1 Model Assessment	Model performance has been assessed using statistically valid/supportable method.	▪ Obtain test strategy document. ▪ If document is not available, interview analytics personnel regarding model testing methods. ▪ Verify that models have been evaluated.
AC 5.2 Model Ranking	Models have been ranked according to performance measures.	▪ Have analytics personnel run models in the presence of internal audit and observe performance statistics. Analytics personnel should explain the nature of model performance measures to internal audit.

Auditing Analytic Models – Controls and Tests

Output Controls (AC 6) — Provide assurance that the model output is presented and used in an appropriate and justified manner to ensure it remains consistent and correct. Includes end-user controls.

Domain	Control	Possible Tests
AC 6.1 Model Results	Model results have been evaluated against business success criteria.	▪ Compare model results with stated goals (see AC 2.1).
AC 6.2 Parsimony	Model results are understandable and justifiable to business owners.	▪ Interview business owners to determine if model results meet expectations and intended use. ▪ Have business owners explain high-level model methodology and rationale for selecting the chosen model among competing models.

Auditing Analytic Models – Controls and Tests

Maintenance Controls (AC 7) — Address the need to re-evaluate and refresh analytic models periodically to ensure they are still relevant in the current environment.

Domain	Control	Possible Tests
AC 7.1 Model Deployment Controls	A model deployment strategy exists and is being followed.	▪ Obtain mode deployment strategy document. ▪ Interview business owners to determine whether deployment strategy is being followed.
AC 7.2 Model Refresh Controls	A process exists to periodically review model accuracy and refresh the selected model when required.	▪ Obtain model refresh strategy document. ▪ If document is not available, interview analytics personnel to determine refresh strategy.

Auditing Analytic Models

Competency Profile and Training Plan

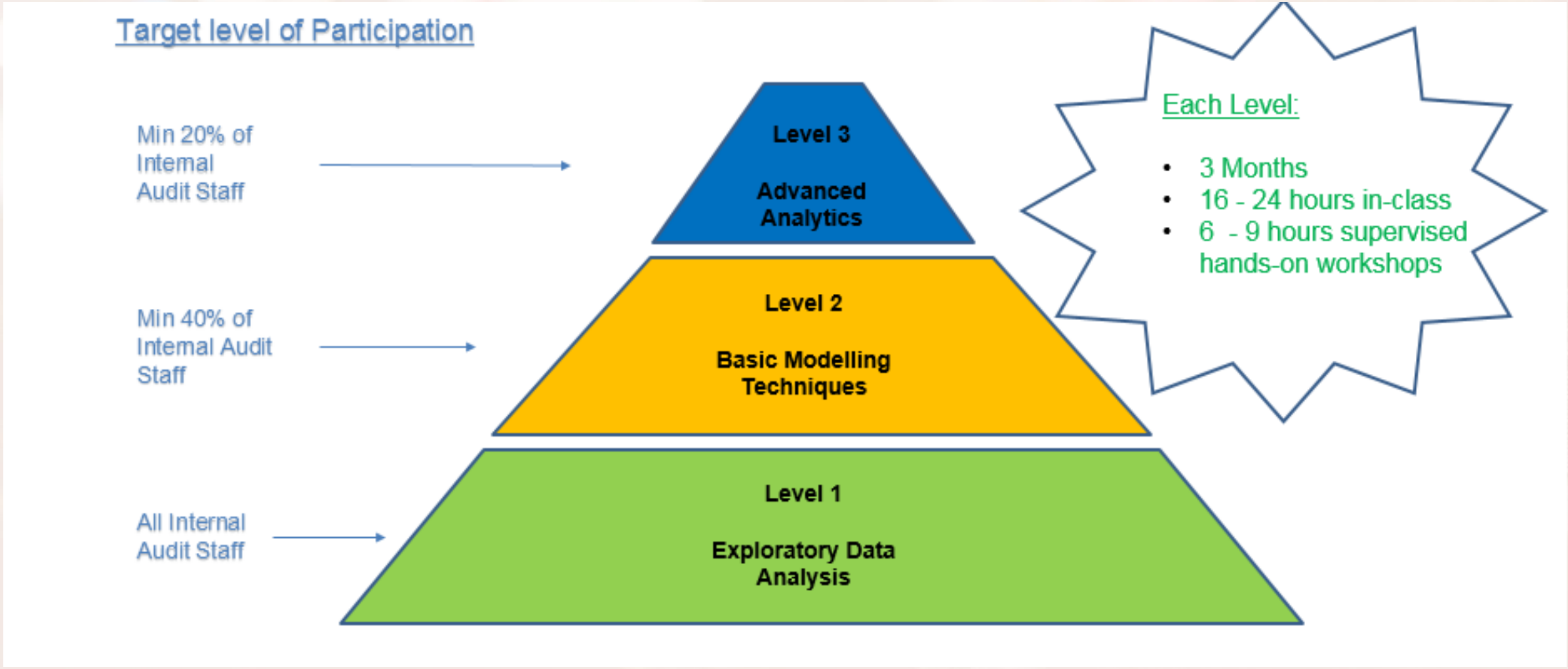
Auditing Analytic Models – IA Analytics Competency Profile

Our organizations will be required to use analytics more and more if they wish to remain relevant and competitive (or remain at all).

Therefore, as Internal Auditors, if we wish to remain relevant, we need to understand and be able to use, audit and consult on a wide variety of analytic tools and techniques:

Analytic Area	Competency – Auditor *(2018)*	Competency – Audit Dept. *(2018)*
Exploratory Data Analysis	Can use 1-2 tools and interpret results, draw inferences.	Can use many tools and interpret results, understands statistical tests, dimensionality reduction strategies, etc.
Basic Modelling Techniques	Understands and can create/use some basic regression/classification models. Linear regression, logistic regression, decision trees.	Understands and can use most basic modelling techniques and some complex methods , clustering, penalized regression, etc., can evaluate model performance.
Advanced Analytics	Understands the principles of machine learning, AI. Can propose potential applications and explain to clients at a conceptual level.	Understands and can create basic prototype ML/AI models. Can explain and discuss with clients at a detailed level.

Auditing Analytic Models – CPC IA Training Plan




fmi
igf
QUESTIONS???