

# Machine Learning and AI for QC of Media Files

Dom Jackson – Telestream LLC

# Overview

- What do we mean by QC?
- What do we mean by Machine Learning and Artificial Intelligence?
- The application of Machine Learning to media QC
- The training problem
- Where we are today
- Where we might go in the future
- Ethics and economics

What do we mean by QC?

# QC can mean a lot of different things...

QC of a media file is the process of verifying that the media file you have is fit for its intended purpose.....

# Types of QC

- File and container checks
- Metadata checks
- Coded bitstream checks
- Video baseband checks
- Audio baseband checks
- Program layout checks

# Manual, Automated and Assisted QC

- Manual QC
  - Driven by a human.
  - Playout devices and analysis tools (vectorscopes, ppm's etc.)
  - Necessary for subjective quality assessment
- Automated QC
  - No human required
  - Automated system makes Pass / Fail decision
- Assisted (Hybrid) QC
  - Automated system highlights possible problem areas to operator
  - Operator makes final decision

# What do we mean by Artificial Intelligence and Machine Learning?

# Artificial Intelligence:

- Any artificial system that appears to behave intelligently in a specific situation
  - The weak AI hypothesis
  - AI doesn't have to be very intelligent
  - What do we actually mean by intelligent?
- The AI effect - "AI is whatever hasn't been done yet"
- Artificial General Intelligence – True AI
  - A system that exhibits intelligence traits and ability to learn similar to those exhibited by “intelligent animals” such as humans
  - What many people mean when they say AI



How does that help me QC Media?

...it doesn't

# What about Machine Learning?

- "A computer program is said to learn from experience  $E$  with respect to some class of tasks  $T$  and performance measure  $P$  if its performance at tasks in  $T$ , as measured by  $P$ , improves with experience  $E$ ." – Tom M. Mitchell
- Automated systems that can learn to perform specific tasks

Sounds like we might be on to  
something here.....

...yes we are

# Why Machine Learning is useful for Media QC (...and many other things)

- ML systems are good at solving types of problems that are difficult to solve by conventional programming methods
- Frequently these are tasks that previously could only be performed by humans.
- ML systems often perform these tasks faster and with fewer errors than humans – machines don't get tired

Many Media QC tasks are amenable to ML approaches:

**A/V Sync**      **Language validation**  
**Set in shot**      **Text positioning**  
**Nudity detection**      **Visual defects**  
**Caption Validation**      **Embargoed words**

This sounds great....

...but there is a problem (nothing is free)

# The Training Problem:

- In order for ML systems to function they need to be trained
- Training can be extremely expensive
- Training typically requires large pre-classified data sets
- Such data sets may require large amounts of human input to curate and classify them

# Mitigating the training problem:

- Use existing data sets
- Generate artificial data sets (....can be risky)
- Leverage existing trained ML systems



# Pre-existing trained ML systems that can be leveraged for Media QC

- Speech to text conversion
- Automated Translation
- Automated Language Recognition
- Text Recognition
- Object recognition
- Inappropriate content recognition

# Most ML is currently in the cloud

- Large cloud providers have the resources to train useful ML systems
- Cloud based ML systems are aimed at an application range far broader than media QC
- Cloud ML benefits from feedback from large user base
- Not always convenient when you are dealing with large media files

So what is out there  
that we can use today?

# Amazon ML Services

- Speech to text
- Text Translation
- Visual Recognition (Rekognition)
  - Object recognition
  - Text recognition
  - Inappropriate content detection

# Google

- Speech to text
- Text Translation
- Visual Recognition
  - Object recognition
  - Inappropriate content detection

# IBM Watson

- Speech to text
- Text Translation
- Visual Recognition

# Microsoft Azure Cognitive Services

- Speech to text
- Text Translation
- Visual Recognition

So where is all of this going?

# The future:

- Current ML systems will continue to improve and add new capability (capability arms race)
- Ability to deploy cloud developed ML models locally
- Increased migration of media workflows to the cloud
- Who knows when “True AI” capabilities will become useful, or where they might take us?

# Ethics and economics

- The cost justification for automation is almost always reducing personnel costs
- This makes ML and other automation less attractive in markets where labor is cheap
- Some media QC problems are so specific to the space that training an ML system may be prohibitively expensive for such a small market



# Conclusions

- Stop calling it AI, and talk about machine learning
- There are a number of existing cloud based ML systems that are useful today for media QC
- Capabilities of ML systems to analyze media will only continue to improve
- ML solutions don't always make economic sense

# Questions?