

Developing **Researching for** Operational AI Weather Service in Taiwan

Hsuan-Tien Lin
林軒田

Professor, National Taiwan University



February 26, 2024
Central Weather Administration

Question from a Meteorologist You-Know-Who (1/5)

Could you provide hints for using GPU supercomputing in developing super-deep DL models?

My Polite Answer

good start with the many GPUs already 😊, what problem do we want to **start solving**?

My Honest Answer

I don't know.

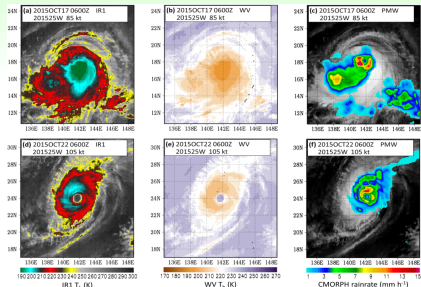
tip: super-deep models come from trying **less deep** ones on **starting problem**

Story 1: Tropical Cyclone Intensity Estimation

meteorologists can 'feel' & estimate TC intensity from image

Can ML/DL do the same/better?

- lack of **ML-ready datasets**
- lack of model that properly utilizes **domain knowledge**



My Honest Thought (2017)

easily solved by a mature Convolutional Neural Network (CNN)?!

no, as explained with two papers

(Chen et al., KDD '18; Chen et al., Weather & Forecasting '19)

From CNN to CNN-TC (Outline)

original CNN **under-estimates** consistently

caused by **dropout** technique within CNN, let's **remove**

CNN without dropout **overfits** more

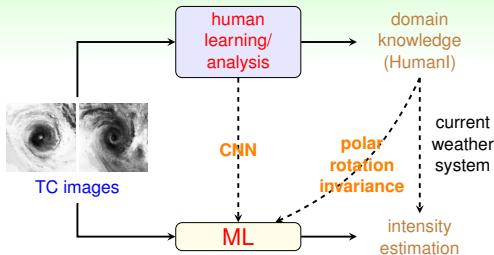
need regularization, leverage **rotation invariance**

wait, TC **rotates differently in different hemispheres!**

ok, let's design model with **polar** rotation invariance

CNN-TC (RMSE 9.03) finally better than
SATCON (9.66) towards **operation-ready**(?)

ML for Modern Artificial Intelligence



Human Learning

- subjective
- produce domain knowledge
- fast basic solution

Machine Learning

- objective
- leverage computing power
- continuous improvement

usually **use humans as much as possible first** before going to machines

Question from a Meteorologist You-Know-Who (2/5)

How can we follow up on rapidly developing DL models and select suitable ones for weather forecasting?

(or similarly) What is the best machine learning model for (my precious big) data and AI?

My Polite Answer

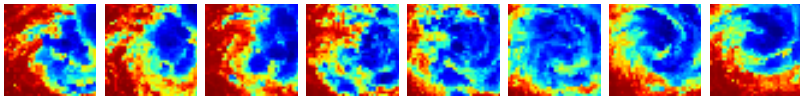
the best model is **problem-dependent**, let's **chat about your problem first**

My Honest Answer

I don't know.

tip: good models come from **interactive modificationS** to meet **problem goal**

Story 2: TC Rapid Intensification Identification



My Honest Thought (2019)

easily solved by a mature Recurrent Neural Network (RNN)?!

no, as explained with (Bai et al., ECML/PKDD '20)

TC Rapid Intensification with Satellite Img. (Selected)

dataset ready? **yes!**

<https://www.csie.ntu.edu.tw/~htlin/program/TCRISI/>

our model is **not much better than competitors**, why?

ah, common evaluation focuses on **Brier score (MSE)**, but not suited for **imbalanced classification**. How about **Heidke Skill Score**?

but that's harder to optimize **from ML perspective**

hmm, how about **area under precision-recall curve (PR-AUC)**?

truth: a paper was written,
but **nobody cares**

Question from a Meteorologist You-Know-Who (3/5)

You have successfully cooperated with a team of meteorologists and forecasters. If you went back in time and did that again, what is the first thing to improve and make it further successful?

evaluation criteria!!



(free image by Manfred Steger from Pixabay)

suggest improvement opportunities



data

hint
preparation steps

techniques

assist
model/tech. choices

usage

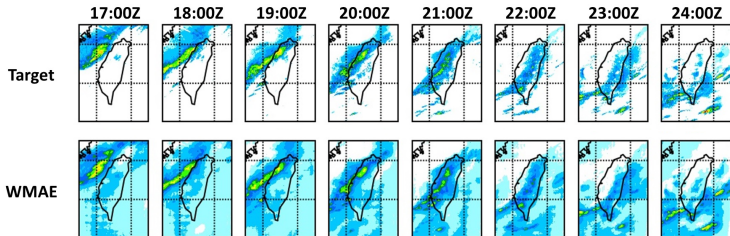
define
acceptance goals

tip: always start with
reasonable & measurable criteria
to describe prioritized **AI goal**

Story 3: Precipitation Nowcasting

My Honest Thought (2020)

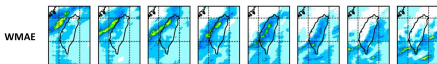
easily solved by a mature Recurrent Neural Network (RNN)?!



no, why is it always raining for our first RNN model? (Ashesh et al., AI for the Earth Systems '22)

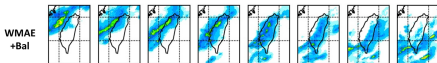
Clear Quantitative Precipitation Nowcasting (Outline)

our model is **always raining**, why?



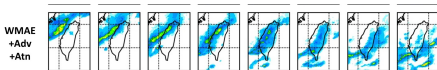
regression model feels **“safer”** to predict a bit of rain

let's force no-rain by **discretizing regression output**



but the rainmap looks **unnatural**

make the rainmap **(visually) indistinguishable to humans**



important to take **iterative professional inputs** from domain experts

Question from a Meteorologist You-Know-Who (4/5)

Suggestion for students and researchers with atmospheric science background to learn deep learning?

For Programming Your Own Idea

nice to have, **but not necessary**

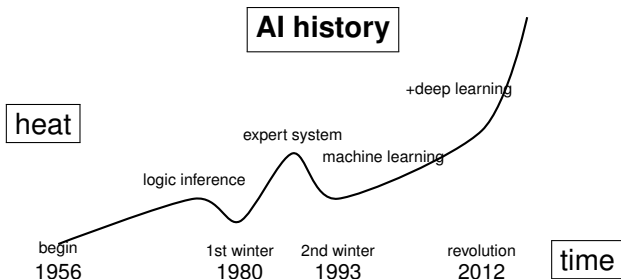
For Communicating with ML/DL Experts

main purpose: knowing what to **expect**

tip: describe your **goal**, understand how ML/DL experts **model** toward your goal, and then provide **feedback**

Question from a Meteorologist You-Know-Who (5/5)

Do you know some cases of failure in inter-discipline AI applications, and what can we learn from them?

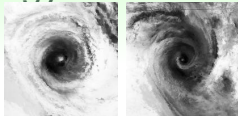


expectation mismatch: the key sin for **application intelligence**

- first AI winter: AI cannot solve 'combinatorial explosion' problems
- second AI winter: expert system failed to scale

tip: make expectation **lower** for **first AI project**

Lessons Learned from Meteorology + ML/DL Research



- 1 yes, **cross-domain collaboration** important
e.g. even from ‘organizing data’ to be ML-ready
- 2 not easy to claim **operation ready**
—can ML be used for ‘**unseenly-strong** TC’?
- 3 successful (operational) AI system requires
 - matching expectations with **clear evaluation criteria**
 - blending human experts with machine with **interactive modifications**
- 4 difficulty: **convince ML/DL researchers** to collaborate
—hope that some of our success stories help

Thank you!