

Нейросетевая революция в метеорологии

А.Б. Колкер , 2024
ФГБУ СибНИГМИ, науч.рук, ктн, НГТУ доц.

В чем разница между искусственным интеллектом и машинным обучением?

Искусственный интеллект (ИИ)
– это общий термин для
различных стратегий и методов,
используемых для того, чтобы
сделать машины более
похожими на людей.

Машинное обучение (МО) – одна из многих других отраслей искусственного интеллекта. Машинное обучение – это наука о разработке алгоритмов и статистических моделей, которые компьютерные системы используют для выполнения сложных задач без четких инструкций. Вместо этого системы полагаются на закономерности и выводы и шаблоны данных.

Нейросеть - это алгоритм (МО) для решения одной из подзадач ИИ



Что такое машинное обучение?

Машинное обучение, представляет собой науку (и искусство) программирования компьютеров для того, чтобы они могли обучаться на основе данных (Орельен Жерон).

Машинное обучение это научная дисциплина, которая наделяет компьютеры способностью учиться, не будучи запрограммированными (Артур Самуэль 1959)

Говорят, что компьютерная программа обучается на основе опыта E по отношению к некоторой задаче T и некоторой оценки эффективности P , если ее эффективность, измеренная на T , измеренная посредством P , улучшается благодаря опыту P (Том Митчел, 1997г).

Области применения МО

Сегментация: “этот кусок поля антициклон, а это - циклон”.

Классификация: “на этом поле есть циклоны и антициклоны, а их положение тут”.

Коррекция: “на этом поле размечен циклон, но его характеристики по-видимому вот такие”.

Прогнозирование: “на следующих временных отсчетах циклон будет находится там”

Типы МО

С учителем (обучающей набор размечен)



Теория и практика

Рисунок Е. ЩЕГЛОВА

Обучение

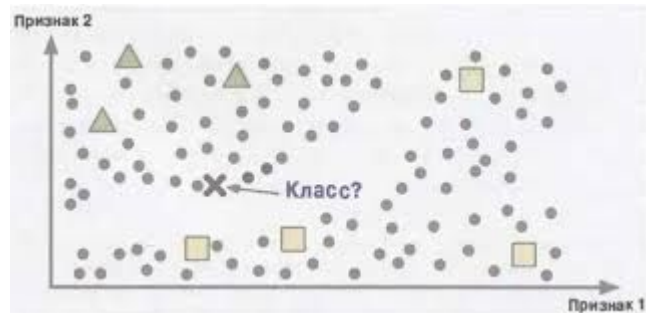
С подкреплением: система тестирует знания и получает за это штрафные или премиальные баллы



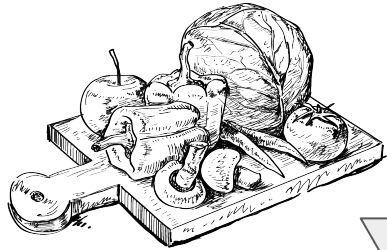
Без учителя: обучающий набор не размечен, система сама ищет закономерности



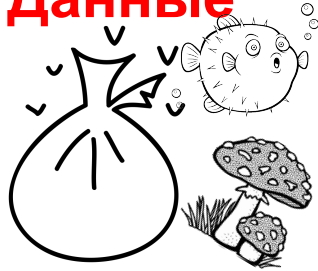
Частичное обучение:
Размечена лишь часть образцов



Машинное обучение - постпроцессинг: как часто его представляют и реализуют ...



Данные



Результат

ЧПП и МО

Методы ЧПП повторяют математически, с упрощениями и допущениями физические законы, по которым работает природа.

Наши знания о физике процессов и возможности математики для их представления непрерывно совершенствуются, однако все детали пока неизвестны и вряд ли будут когда-либо полностью воспроизведены.

Методы NWP хорошо масштабируются с вычислительными ресурсами: точность обычно возрастает с увеличением вычислительных ресурсов, поскольку дополнительные вычислительные ресурсы могут быть выделены на моделирование с более высоким разрешением, с более сложными схемами параметризации (Bauer et al, 2015; Benjamin et al., 2019)

Основной способ совершенствования методов NWP - это ручная разработка высококвалифицированными экспертами более совершенных моделей, алгоритмов и аппроксимации, что является трудоемким и дорогостоящим процессом.

Методы ML пытаются подстроиться под физику без понимания ее. В широком спектре научных областей точность методов на основе ML может повышаться по мере поступления все большего объема данных более высокого качества, часто при гораздо меньшем вычислительном бюджете. В прогнозировании погоды системы ML начинают превосходить модели прогнозирования на основе NWP. В качестве примера можно привести субсезонные прогнозы тепловых волн (Lopez-Gomez et al., 2022) и прогнозирование осадков как с использованием информации ЧПП, так и без нее (Espeholt et al., 2022; Ravuri et al., 2021; Shi et al., 2017; Sønderby et al., 2020).

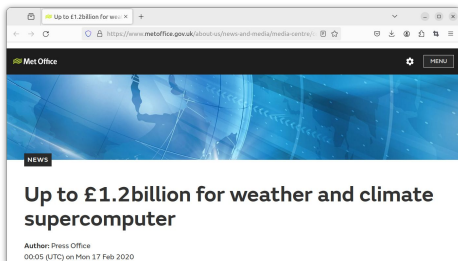
ЧПП и МО

улучшить ЧПП:

- изучить еще более тонкие детали: как работает физика;
- описать их математически;
- запрограммировать алгоритм;
- построить еще более мощный компьютер (*увеличение мощности вычислителя один из наиболее значимых факторов улучшения прогноза*);

Компьютер для ЧПП - десятки ... сотни миллионов долларов

История ~ 70 лет развития



улучшить МО:

- улучшить архитектуру алгоритма МО;
- задействовать большой объем качественных данных;
- построить графический ускоритель соответствующей мощности (*увеличение мощности вычислителя полезно, не является главным фактором улучшения прогноза*).



Компьютер для МО - десятки... сотни тысяч долларов (мин требования для обучения: GPU 32GB)

История- 10-15 лет развития, первые осязаемые результаты - 2022 (Keisler 2022, s. Kurth et al. 2022; Pathak et al. 2022 FourCastNet(FCN),Remi Lam GraphCast 2022)



Плата для ускорения вычислений NVIDIA Tesla V100 PCIe 5120
cores 4096-bit 32 GB CoWoS HBM2 with ECC
Цена: 600 - 800 тыс руб., облако около 18 USD\час

История вопроса

1943г «Логическое исчисление идей, относящихся к нервной активности»¹ Уоррен Мак-Каллок и Уолтер Питтс - понятие искусственной нейронной сети[1]

1949г «Организация поведения» д.Хебб - процесс обучения нейронной сети

1957г Перцептрон Френка Розенюлатта

.....

1980 Кунихо Фукусима - многослойная нейронная сеть и обратное распространение ошибки

1997 Длинная цепь элементов краткосрочной памяти (Long short-term memory, LSTM)

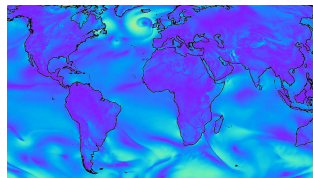
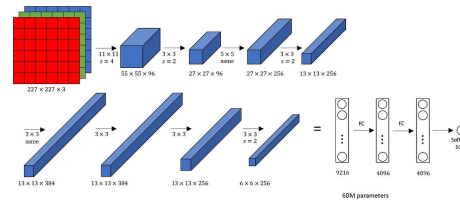
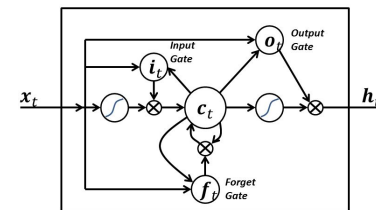
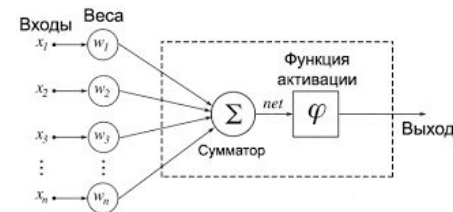
(S Hochreiter 1997)]

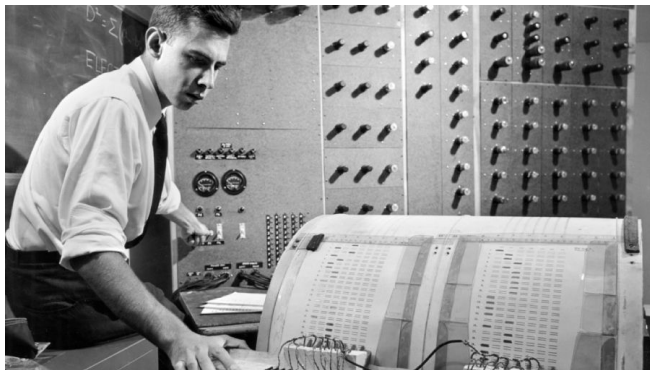
1998 году Ян Лекун создал сверточную нейронную сеть

2012 г AlexNet Alex Krizhevsky University of Toronto

2022 Языковые модели

2023 GraphCast и др





Фрэнк Розенблатт (US)
(1928-1971)



Кунихико Фукусима (JP) – создатель многослойной нейронной сети
(Fuzzy Logic Systems Institute in Fukuoka, Japan)



Ян Лекун (FR) - многослойная сверточная сеть (META)



Remi Lam (UK) Google Deep Mind , MIT

Что у нас?

Гидрометцентр РФ:

- 2000е Схема РЭП.
- 2014 А. Багров - “Комплексный прогноз приземных метеорологических величин”.
- 2022 Ф.Быков (Постпроцессинг численных прогнозов приземных метеорологических параметров на основе нейросетевых методов).
- М. Цырульников - исследования в области ассимиляции данных

ФГБУ СибНИГМИ

- Здерева М.Я - методы постпроцессинга на базе деревьев

НИЦ Планета

- системы постобработки ДЗЗ.

... Кого-нибудь обязательно забыл (заранее простите, не специально)

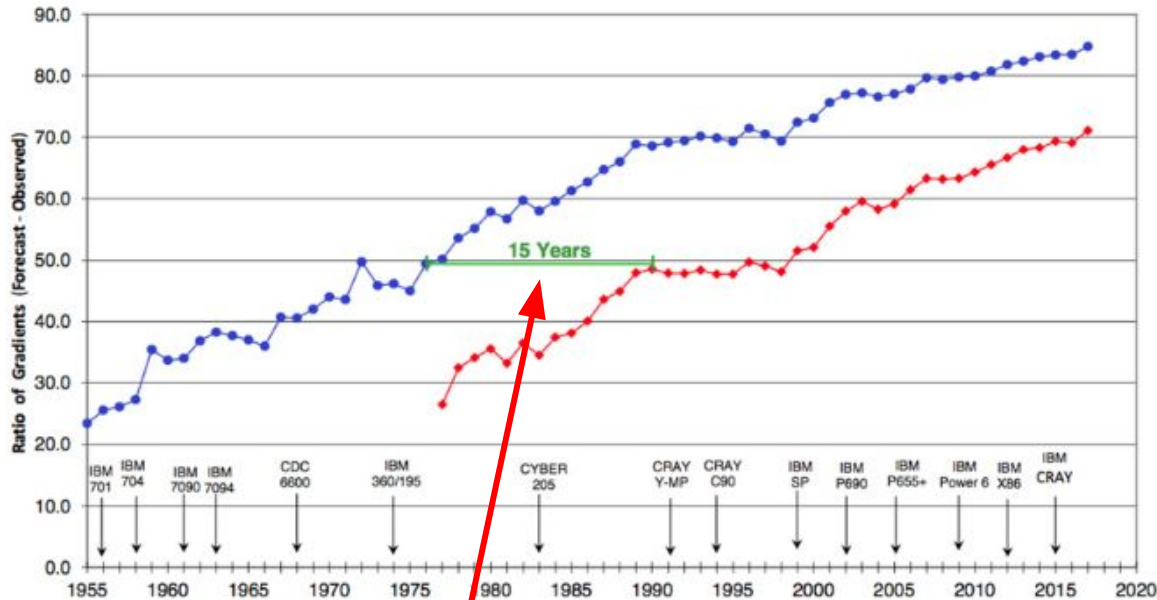


NCEP Operational Forecast Skill

36 and 72 Hour Forecasts @ 500 MB over North America [100 * (1-S1/70) Method]



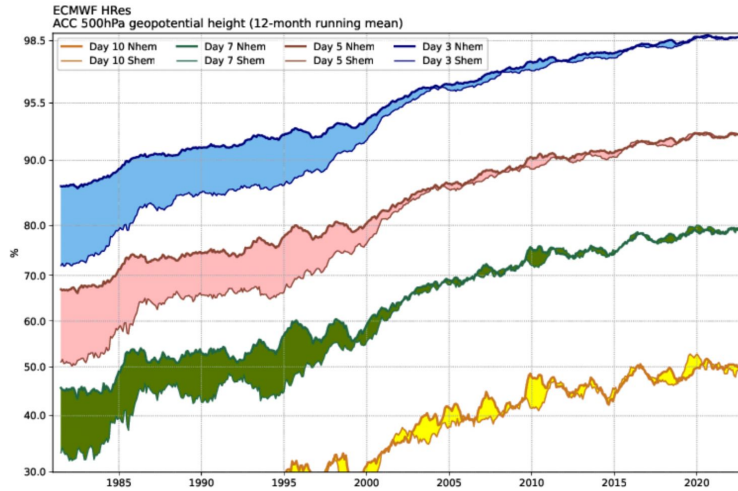
—●— 36 Hour Forecast —●— 72 Hour Forecast



Benjamin, Stanley & Brown, John & Brunet, Gilbert & Lynch, Peter & Saito, Kazuo & Schlatter, Thomas. (2019). 100 Years of Progress in Forecasting and NWP Applications. Meteorological Monographs. 59. 10.1175/AMSMONOGRAPH S-D-18-0020.1.

Качество прогноза на 72 часа отстает от прогноза на 36 часов на 15 лет исследований. S1 score - метрика, которая (упрощенно) отражает насколько хорошо прогностическая модель прогнозирует аномалии

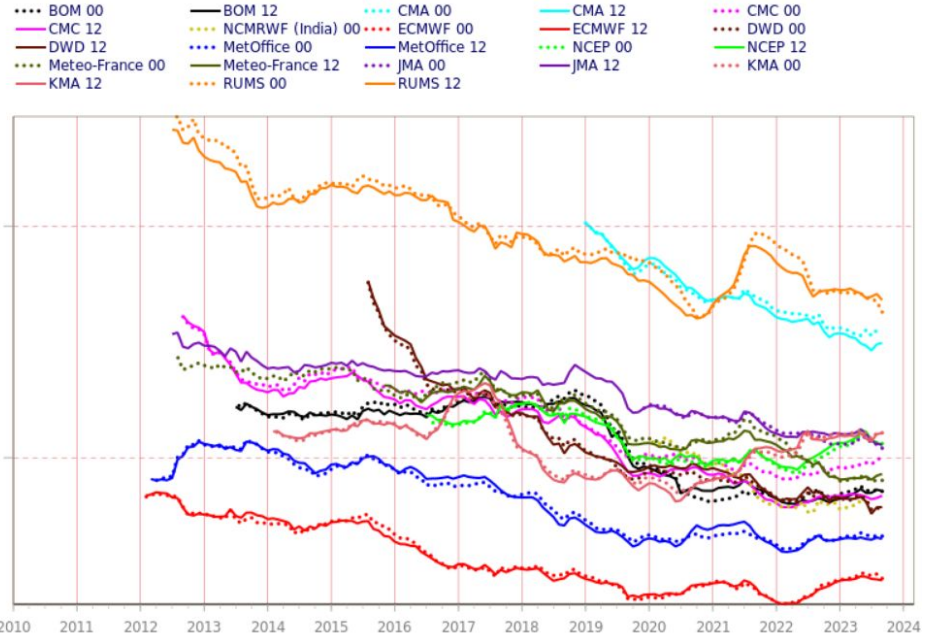
Lead time of anomaly correlation coefficient (ACC) reaching multiple thresholds (High resolution (HRES) 500 hPa height forecasts)



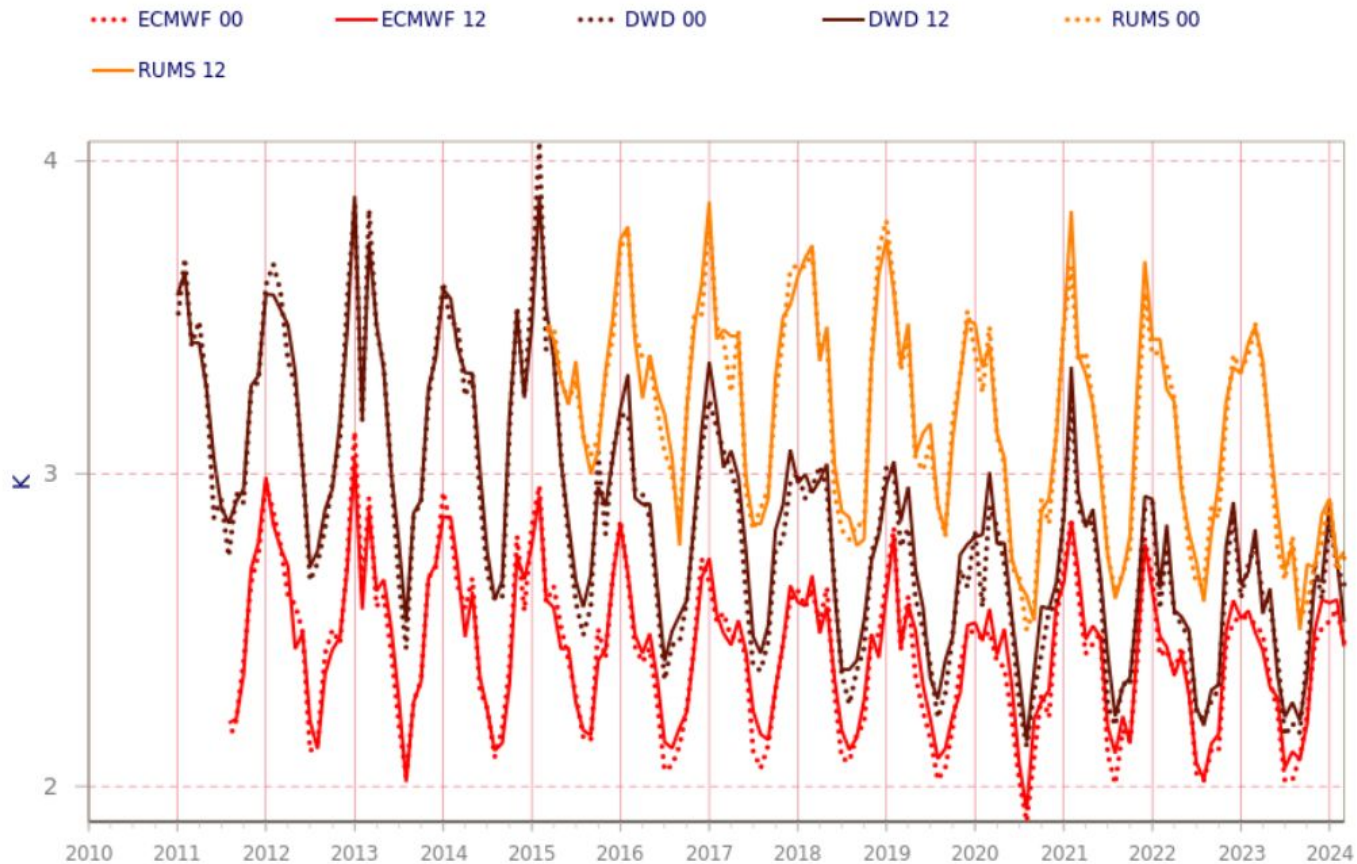
© 2023 European Centre for Medium-Range Weather Forecasts (ECMWF)
Source: www.ecmwf.int
License: CC BY 4.0 and ECMWF Terms of Use (https://apps.ecmwf.int/datasets/#/concepts/general)
Created at 2023-11-10T22:05:41.04Z



Step: 96 RMSEF/msl/n.hem/analysis mean 12m



Step: 120 RMSEF 850 hPa t/n.hem/observations



И тут в аптеке появились кузнецы(*)...

- **FourCastNetv2:** NVIDIA 0.25`
- **Graphcast:** GoogleDeepMind 0.25`
- **Pangu-Weather:** Huawei 0.25`
- **AIFS:** - ECMWF 1`, 13 уровней

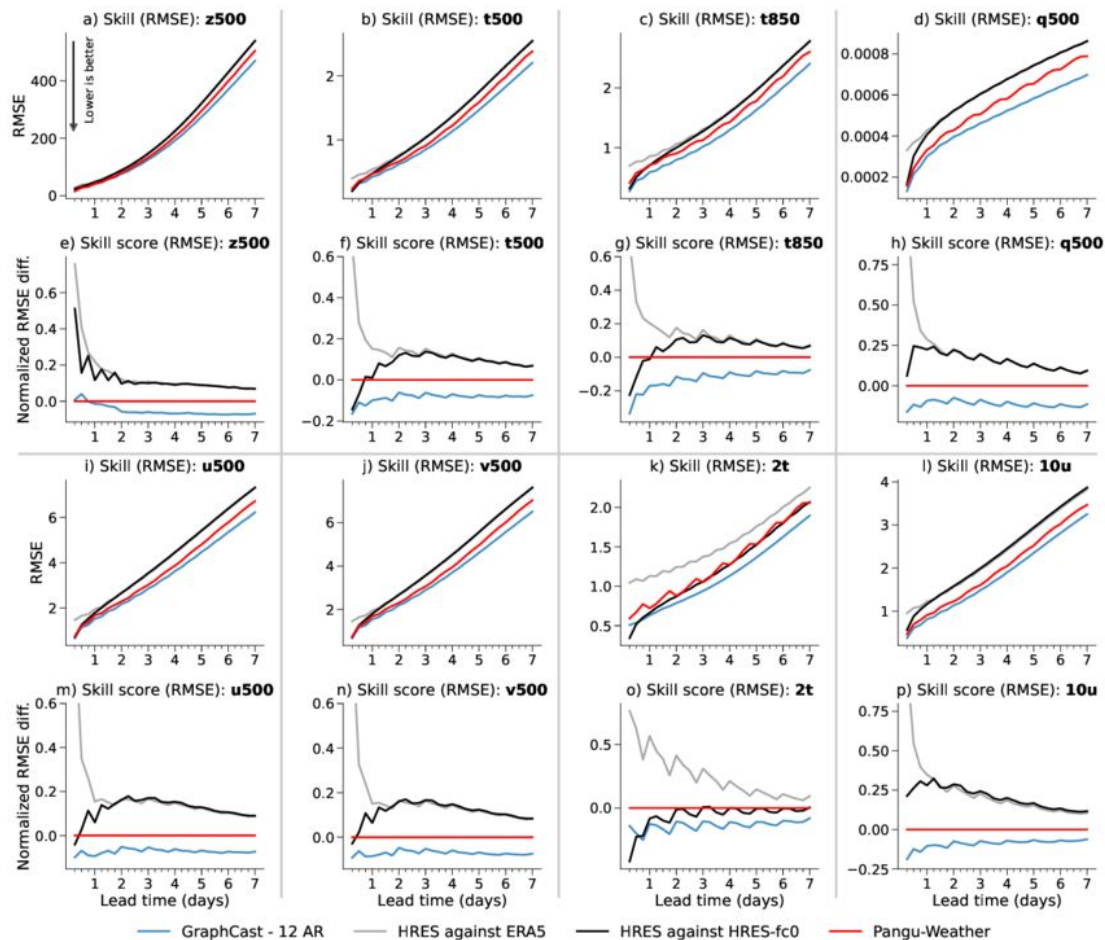


Из 4х перечисленных моделей, только AIFS разрабатывается в центре, который исторически занимается погодой!



(*) на самом деле кузнецы были математиками и программистами

И получилось у них ни разу не хуже



Сравнение GraphCast и Pangu-Weather по показателю RMSE. Строки 1 и 3 показывают абсолютный RMSE для GraphCast (синие линии), Pangu-Weather (Bi et al., 2022) (красные линии), HRES, оцененный по HRES-fc0 (черные линии), и HRES, оцененный по ERA5; строки 2 и 4 показывают нормализованные различия RMSE между моделями по отношению к Pangu-Weather. Ось x представляет шаг по времени с приращением 6 часов в течение 10 дней. Нижняя строка - нормализованный RMSE, относительно Pangu-Weather

Основная статья GrapCast

GraphCast: Learning skillful medium-range global weather forecasting

Remi Lam*,¹, Alvaro Sanchez-Gonzalez*,¹, Matthew Willson*,¹, Peter Wirnsberger*,¹, Meire Fortunato*,¹, Alexander Pritzel*,¹, Suman Ravuri¹, Timo Ewalds¹, Ferran Alet¹, Zach Eaton-Rosen¹, Weihua Hu¹, Alexander Merose², Stephan Hoyer², George Holland¹, Jacklynn Stott¹, Oriol Vinyals¹, Shakir Mohamed¹ and Peter Battaglia¹ *equal contribution, ¹DeepMind, ²Google



Remi Lam (UK) Google Deep Mind , MIT

Где революция?

Традиционным гидродинамическим подходам потребовались десятилетия, чтобы достичь таких показателей, которые нейросетевые модели погоды показали на прямо старте. *

* Однако тут есть некоторая доля лукавства- для расчета нейросетевой модели (пока) требуются начальные поля - несколько последовательных полей реанализа или ассимиляции (например, ERA5 или IFS), которые есть результат тех самых десятилетий совершенствования ГДМ и ЧПП.

Как это работает? (упрощенно для понимания)

Поля значений ->
свертка -> общая
картина $t, d-dt$



Общая картина, ее
динамика - граф-тензор

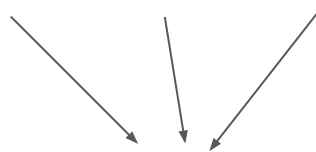


Декомпозиция в поле
значений $t + dt$

227 переменных (батм x 37
уровней + 5 на поверхности за
два срока)

Обучение:

40 лет ECMWF ERA5

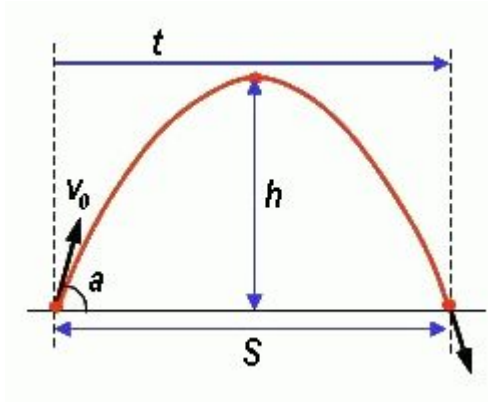
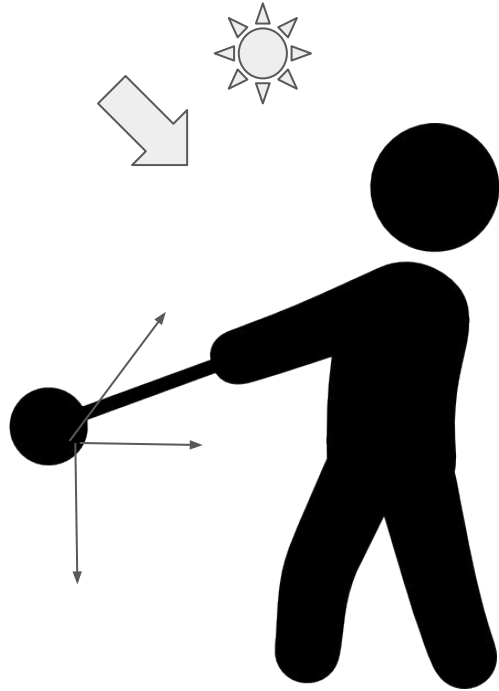


Энкодер

Процессор

Декодер

Представим метателя молота



$$S = \frac{v_0^2}{g} \sin 2\alpha$$

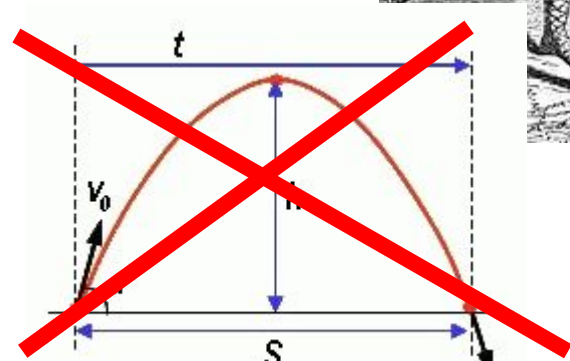
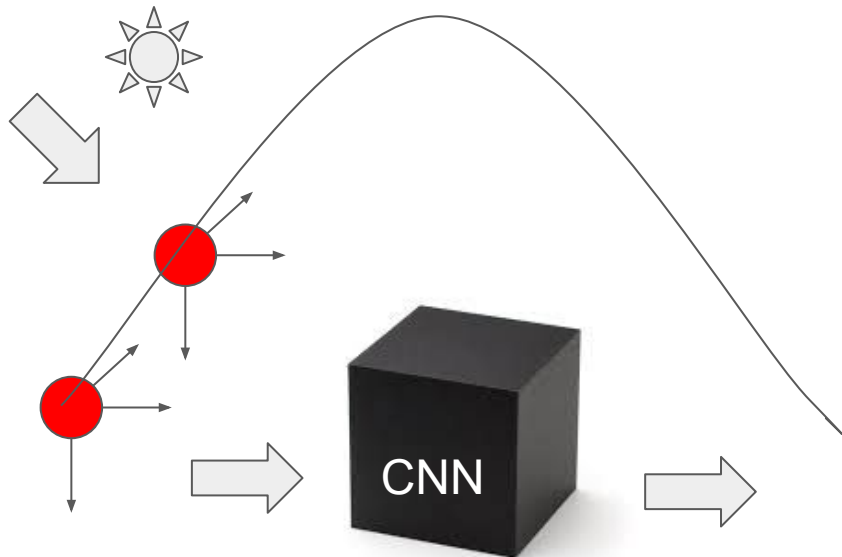
$$h = \frac{v_0^2}{2g} \sin^2 \alpha \quad t = \frac{2v_0 \sin \alpha}{g}$$



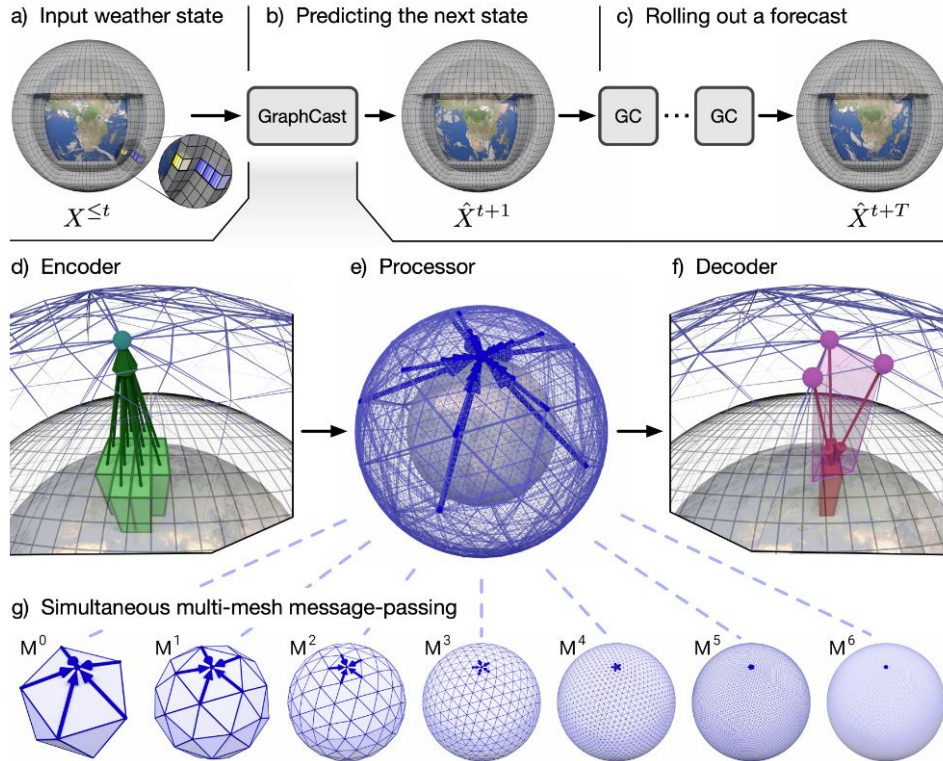
- v, α, g
- t, p, v_u, v_v



Мы понятия не имеем про физику, зато наблюдаем за метателями молота с античности и знаем куда он обычно летает в тех или иных условиях



Как это устроено (на примере GraphCast)?



(R.Lam at all 2022)

5 приземных, 6 атм на 37
уровнях = 227 всего.
 $0.25^\circ : 1 \square 038 \square 240 = 721$
1440

Графовая нейронная сеть:
подобно ячейке ЧПП, есть
состояние(узел и его
состояние по вертикали) и
есть обмен “энергией” с
соседними ячейками(ребра) и
“правила”, по которым
выполняется этот обмен

ВХОД-ВЫХОД

Type	Variable name	Short name	ECMWF Parameter ID	Role (accumulation period, if applicable)
Atmospheric	Geopotential	z	129	Input/Predicted
Atmospheric	Specific humidity	q	133	Input/Predicted
Atmospheric	Temperature	t	130	Input/Predicted
Atmospheric	U component of wind	u	131	Input/Predicted
Atmospheric	V component of wind	v	132	Input/Predicted
Atmospheric	Vertical velocity	w	135	Input/Predicted
Single	2 metre temperature	2t	167	Input/Predicted
Single	10 metre u wind component	10u	165	Input/Predicted
Single	10 metre v wind component	10v	166	Input/Predicted
Single	Mean sea level pressure	msl	151	Input/Predicted
Single	Total precipitation	tp	28	Input/Predicted (6h)
Single	TOA incident solar radiation	tisr	212	Input (1h)
Static	Geopotential at surface	z	129	Input
Static	Land-sea mask	lsm	172	Input
Static	Latitude	n/a	n/a	Input
Static	Longitude	n/a	n/a	Input
Clock	Local time of day	n/a	n/a	Input
Clock	Elapsed year progress	n/a	n/a	Input

Наилучшее разрешение по времени 1ч до 90 ч, 3ч 93-144, 6ч 150-240ч (Rasp et al., 2020).

Сравнение свойств

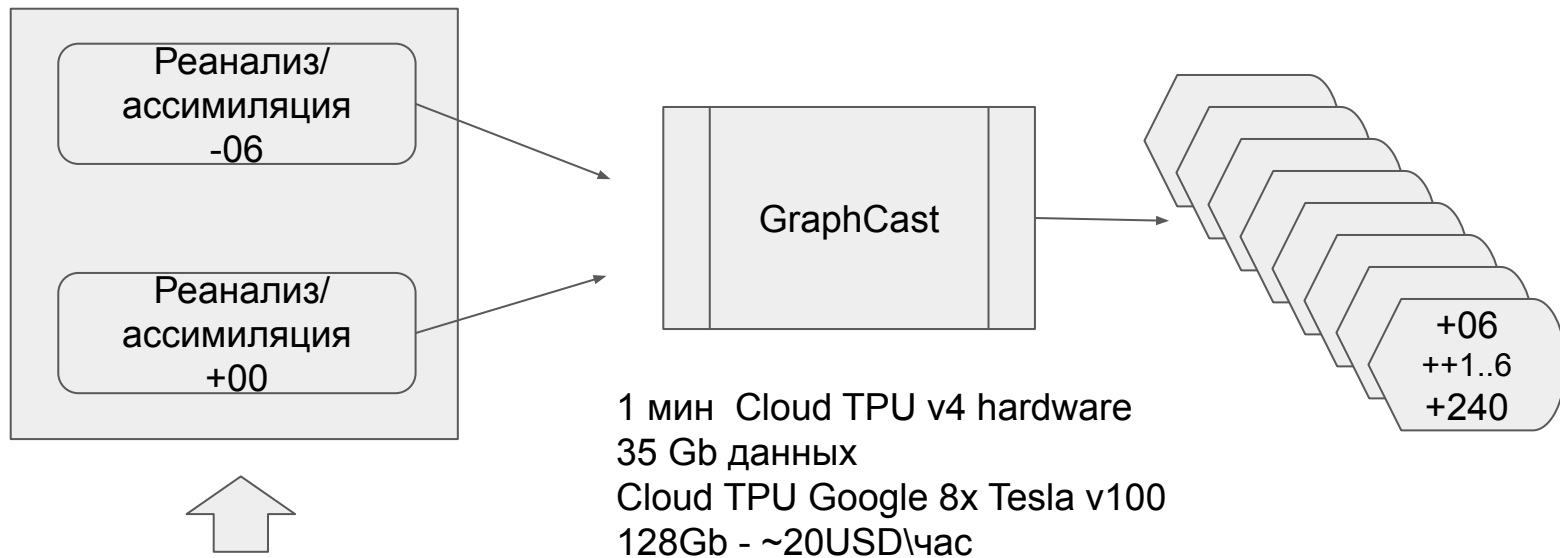
NWP:

- большое количество параметров и уровней;
- шаг продукции по времени и пространству настраивается;
- “механика” основана на физике и не может “галлюцинировать” ни при каких обстоятельствах
- выходные параметры могут быть расширены путем расширения физических блоков без перенастройки всей модели;
- расчет- долго, дорого, и ресурсоемко.

MO

- уровни и параметры фиксированы;
- шаг продукции по времени фиксирован (6ч Graphcast);
- любое изменение в модели = полный процесс переобучения (намного дольше и энергозатратнее, чем сам процесс создания прогноза)
- механика не основана на физике, вероятность “галлюцинации” не равна нулю.
- обучение долго дорого и ресурсоемко, но создание продукции- быстро

Нейросетевые модели (пока*) полностью зависят от традиционных ЧПП

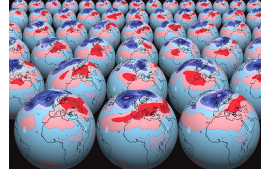


физика + математика = поля
плюс тензор динамики -
традиционные ЧПП

(*)Peter Battaglia, Google DeepMind - мы работаем над
стартом модели непосредственно от измерений без усвоения

(GraphCast: Learning skillful medium-range global weather forecasting Dr Peter Battaglia, DeepMind

Friday 12th May 2023)



Направления развития: гибрид, ансамбль, ...

- прогноз ОЯ НЯ
- ансамбли

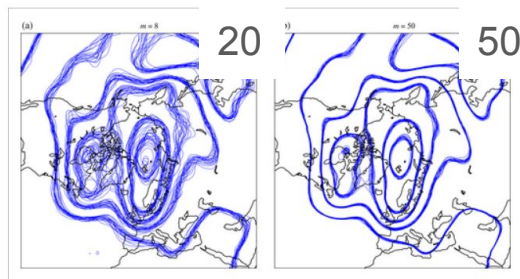
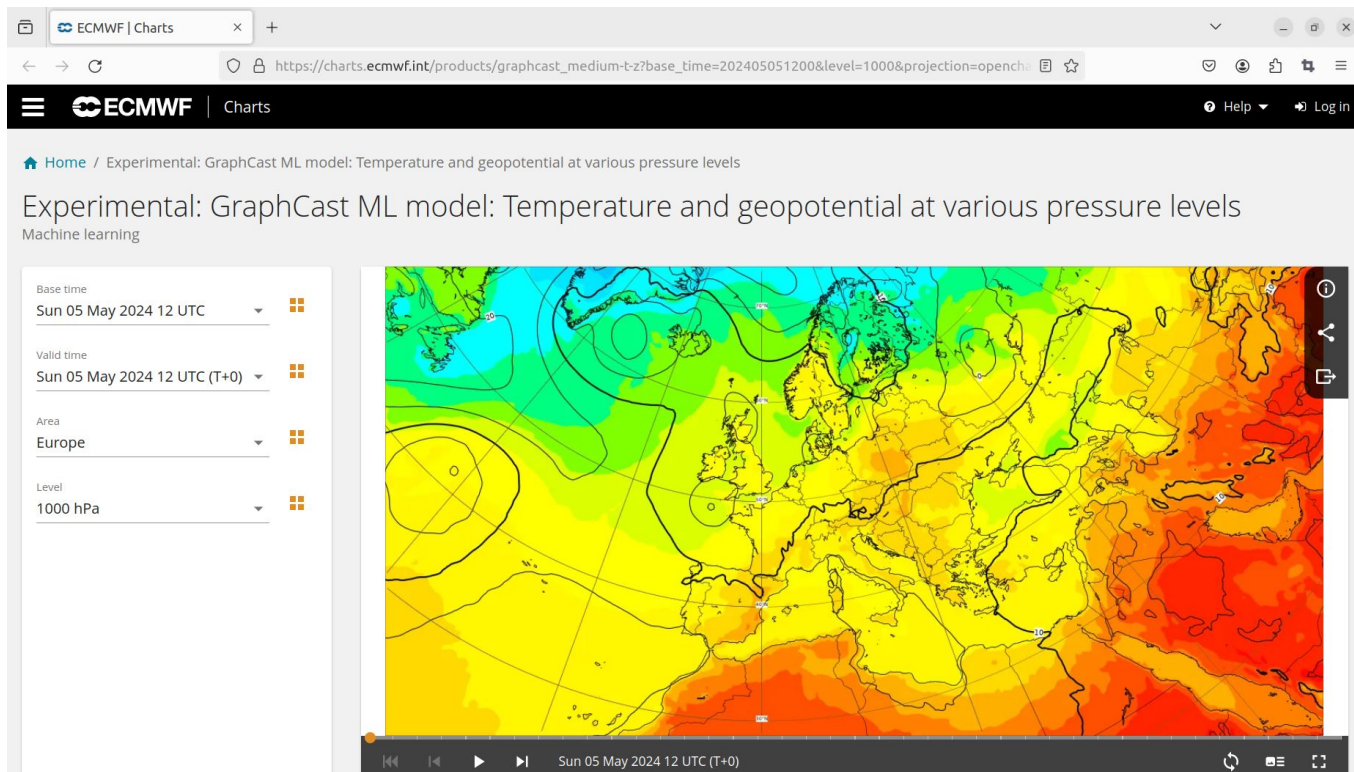


Рис. (Leutbecher M. 2019)

Organization	Country	M
CMA	China	14
CPTEC	Brazil	14
ECCC	Canada	20
ECMWF	International	50
JMA	Japan	26
KMA	Korea	24
Meteo-France	France	34
MetOffice	United Kingdom	17
NCEP	USA	20
NCMRWF	India	44
EN_SLaV	ГМЦ (2022)	40

-
-
- старт ЧПП от расчетов МО (пока не хватает переменных)
- коррекция полей ЧПП по данным нейросетевых моделей.

Результаты моделирования уже доступны в открытом доступе





ECMWF | Charts

https://charts.ecmwf.int/?facets={"Product type"%3A["Experimental%3A Machine learning models"]}

ECMWF | Charts

Home / Charts catalogue

Search products...

Range

- Medium (15 days)
- Extended (42 days)
- Long (Months)

Type

- Forecasts
- Verification

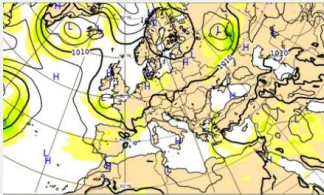
Component

- Surface
- Atmosphere

Product type

- High resolution forecast (HRES)
- Ensemble forecast (ENS)

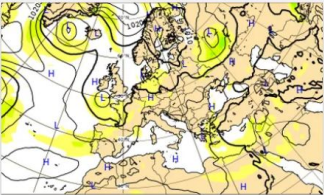
Latest forecast



Experimental: AIFS (ECMWF) ML model: Mean sea level pressure and 850 hPa wind speed

AIFS (ECMWF): a deep learning-based system developed by ECMWF. It is initialised with ECMWF HRES analysis. AIFS operates at 0.25° resolution

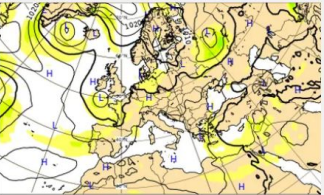
Latest forecast



Experimental: FourCastNet ML model: Mean sea level pressure and 850 hPa wind speed

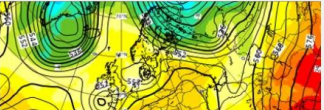
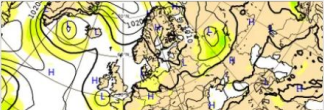
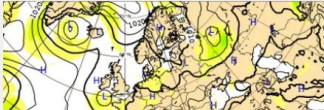
FourCastNet v2-small: a deep learning-based system developed by NVIDIA in collaboration with researchers at several US universities. It is initialised with ECMWF HRES analysis. FourCastNet operates at 0.25° resolution.

Latest forecast



Experimental: FuXi ML model: Mean sea level pressure and 850 hPa wind speed

FuXi: a deep learning-based system developed by researchers at Fudan University. It is initialised with ECMWF HRES analysis. FuXi operates at 0.25deg resolution.



Итог

- Технологии постпроцессинга расчетов ЧПП на базе машинного обучения успешно применяются уже течения десятилетия.
- Технологии “прямого” прогноза на базе нейросетевых решений - новое слово в синоптическом ремесле. Очевидно, что в ближайшее время их ожидает бурное развитие, но его темп трудно прогнозируем.
- Нейросетевой прогноз уже сейчас оказывается успешнее, чем традиционные ЧПП, особенно на больших заблаговременностях (T,P,H,UV). Данных, чтобы сделать вывод по осадкам пока недостаточно.
- Нейросетевой прогноз (пока) не может существовать без ЧПП (усвоение)
- Нейросетевой прогноз (пока) уступает традиционным ЧПП по шагу сетки (0.25) и перечню параметров.