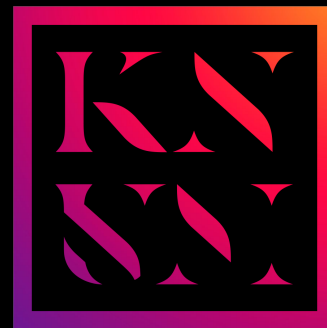


Prezentacja na CORE: co i jak?

Julia Bazińska

Koło Naukowe Uczenia Maszynowego UW

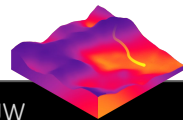
5 listopada 2020





Agenda

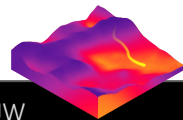
1. Wybór tematu
2. Jak czytać paper?
3. Struktura prezentacji
4. Garść różnych rad





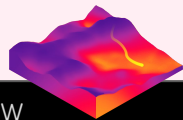
Jaki wybrać temat?

- Twój projekt na studia / w pracy
- Coś, co zawsze chciałaś_eś zrozumieć, ale brakło motywacji
- Można zapytać się kogoś z Zarządu albo na grupie członków






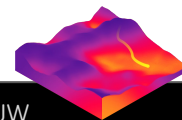
Jak czytać paper?





Jakie byłoby przyzwoite tłumaczenie słowa
"paper"?

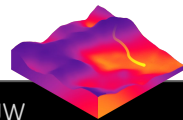
 Start presenting to display the poll results on this slide.



Jak czytać paper?

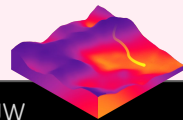
Zadaj sobie pytania:

- jakie zadanie jest rozwiązywane?
 - Klasyfikacja obrazu, rozpoznawanie mowy, model języka, ...
- co jest główną innowacją w paperze?
- co znaczą metryki?
- w czym nowa metoda jest lepsza?
- ...a w czym jest gorsza?





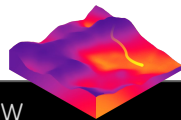
Struktura prezentacji





Przykładowy plan prezentacji

- 1) Motywacja
 - 2) Plan prezentacji
 - 3) Related work
 - 4) Nowa metoda wprowadzona w paperze
 - 5) Przeprowadzone eksperymenty
 - 6) Wyniki + użyte metryki
 - 7) Główne wnioski (key takeaways)
- Diagram illustrating the structure of a presentation plan:
- Wstęp (Introduction) covers items 1 and 2.
 - Główny content (Main content) covers items 3, 4, 5, and 6.
 - Zakończenie (Conclusion) covers item 7.

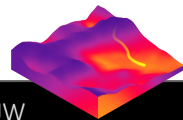




Przykładowy plan prezentacji

- 1) Motywacja
 - 2) Plan prezentacji
 - 3) Related work
 - 4) Nowa metoda wprowadzona w paperze
 - 5) Przeprowadzone eksperymenty
 - 6) Wyniki + użyte metryki
 - 7) Główne wnioski (key takeaways)
- Diagram illustrating the structure of a presentation plan:
- Wstęp (Start)
 - 1) Motywacja
 - 2) Plan prezentacji
 - Główny content (Main content)
 - 3) Related work
 - 4) Nowa metoda wprowadzona w paperze
 - 5) Przeprowadzone eksperymenty
 - 6) Wyniki + użyte metryki
 - Zakończenie (End)
 - 7) Główne wnioski (key takeaways)

...ale to oczywiście zależy, o czym mówicie!

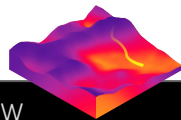




Przykładowy plan -- motywacja

- Cel: zainteresować słuchaczy
- Trochę backgroundu
 - “Przez długi czas do rozpoznawania mowy używano złożonych modeli...”
- Trochę o paperze
 - “Ale okazuje się że można do tego użyć sieci neuronowych, co jest wygodniejsze, bo...”
- Na tyle, żeby słuchacze zrozumieli ogólną ideę

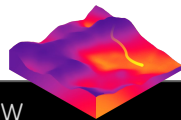
- 1) **Motywacja**
- 2) Plan prezentacji
- 3) Related work
- 4) Nowa metoda wprowadzona w paperze
- 5) Przeprowadzone eksperymenty
- 6) Wyniki + użyte metryki
- 7) Główne wnioski (key takeaways)





Przykładowy plan -- plan prezentacji

- Cel: żeby słuchacze lepiej się orientowali
 - Jak ktoś na chwilę się rozproszy, to łatwiej mu będzie wrócić!
 - Zaznaczenie przejść do nowych sekcji
- 1) Motywacja
 - 2) **Plan prezentacji**
 - 3) Related work
 - 4) Nowa metoda wprowadzona w paperze
 - 5) Przeprowadzone eksperymenty
 - 6) Wyniki + użyte metryki
 - 7) Główne wnioski (key takeaways)



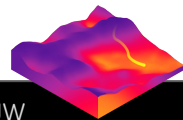


Przykładowy plan -- related work

- Cel: by było wiadomo do czego porównujemy
- jakie jest SOTA? jakie są osiągnięcia w tym zadaniu dotychczas?
 - “W SGD był taki a taki problem, został rozwiązany przez Adagrad, ale...”
 - “Istnieją narzędzia nieoparte o ML, które to robią.”
- może być bardzo krótko
- można w ogóle pominąć jeśli się pokrywa z motywacją

- 1) Motywacja
- 2) Plan prezentacji
- 3) **Related work**
- 4) Nowa metoda wprowadzona w paperze
- 5) Przeprowadzone eksperymenty
- 6) Wyniki + użyte metryki
- 7) Główne wnioski (key takeaways)

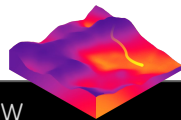
SOTA -- state of the art,
tzn. model z najlepszym
wynikiem w danym zadaniu





Przykładowy plan -- opis innowacji

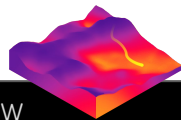
- Cel: wyjaśnić dokładnie, co jest głównym celem papera
 - bardzo zależy od kontentu
 - “nowa warstwa, która działa tak...”
 - “porównanie różnych modeli rekurencyjnych”
 - “nowa metoda augmentacji danych głosowych”
- 1) Motywacja
 - 2) Plan prezentacji
 - 3) Related work
 - 4) Nowa metoda wprowadzona w paperze**
 - 5) Przeprowadzone eksperymenty
 - 6) Wyniki + użyte metryki
 - 7) Główne wnioski (key takeaways)





Przykładowy plan -- eksperymenty

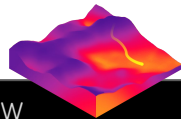
- Cel: wyjaśnić, jak mierzono jakość modelu i jakie ma wyniki
 - “użyto datasetu X, funkcją straty było Z, wynik był taki a taki”
 - te dwie sekcje mogą się łączyć
 - warto wyjaśnić dobrze metryki!
- 1) Motywacja
 - 2) Plan prezentacji
 - 3) Related work
 - 4) Nowa metoda wprowadzona w paperze
 - 5) **Przeprowadzone eksperymenty**
 - 6) **Wyniki + użyte metryki**
 - 7) Główne wnioski (key takeaways)





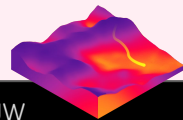
Przykładowy plan -- wnioski

- Cel: podsumowanie!
 - “Zatem mamy metodę X, która przyspiesza uczenie, ale nie pogarsza wyników”
 - “Wyniki są dobre, ale jeszcze daleko od tego, by zastosować ten model w rzeczywistości”
- Powtórzyć parę najważniejszych elementów (dla zagubionych)
 - 1) Motywacja
 - 2) Plan prezentacji
 - 3) Related work
 - 4) Nowa metoda wprowadzona w paperze
 - 5) Przeprowadzone eksperymenty
 - 6) Wyniki + użyte metryki
 - 7) **Wnioski (key takeaways)**





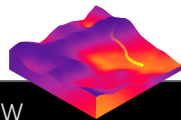
Garść różnych rad





Zaangażuj słuchaczy!

- “jak myślicie, jakie mogą być wady tego podejścia? a jakie mogą być zalety?”
- może być trudne do wplecenia w prezentację
- narzędzia takie jak mentimeter / slido
- dyskusja -> oznacza zrozumienie i zainteresowanie





Wzory matematyczne

$$P(X) = \int P(X|z; \theta) P(z) dz.$$

X - data we're modeling (like images)
z - latent variable (like which number to draw)
 θ - model parameters (network weights)

- opisy oznaczeń!
- skomplikowane przekształcenia nie są istotne
- jeśli wzór ma nazwę, to warto ją podać

Ładny przykład dobrze opisanego wzoru.

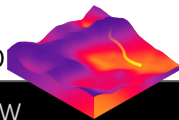




Tabela z wynikami

Method	Depth	Params	C10	C10+	C100	C100+	SVHN
Network in Network [22]	-	-	10.41	8.81	35.68	-	2.35
All-CNN [32]	-	-	9.08	7.25	-	33.71	-
Deeply Supervised Net [20]	-	-	9.69	7.97	-	34.57	1.92
Highway Network [34]	-	-	-	7.72	-	32.39	-
FractalNet [17]	21	38.6M	10.18	5.22	35.34	23.30	2.01
with Dropout/Drop-path	21	38.6M	7.33	4.60	28.20	23.73	1.87
ResNet [11]	110	1.7M	-	6.61	-	-	-
ResNet (reported by [13])	110	1.7M	13.63	6.41	44.74	27.22	2.01
ResNet with Stochastic Depth [13]	110	1.7M	11.66	5.23	37.80	24.58	1.75
	1202	10.2M	-	4.91	-	-	-
Wide ResNet [42]	16	11.0M	-	4.81	-	22.07	-
	28	36.5M	-	4.17	-	20.50	-
	16	2.7M	-	-	-	-	1.64
ResNet (pre-activation) [12]	164	1.7M	11.26*	5.46	35.58*	24.33	-
	1001	10.2M	10.56*	4.62	33.47*	22.71	-
DenseNet ($k = 12$)	40	1.0M	7.00	5.24	27.55	24.42	1.79
DenseNet ($k = 12$)	100	7.0M	5.77	4.10	23.79	20.20	1.67
DenseNet ($k = 24$)	100	27.2M	5.83	3.74	23.42	19.25	1.59
DenseNet-BC ($k = 12$)	100	0.8M	5.92	4.51	24.15	22.27	1.76
DenseNet-BC ($k = 24$)	250	15.3M	5.19	3.62	19.64	17.60	1.74
DenseNet-BC ($k = 40$)	190	25.6M	-	3.46	-	17.18	-

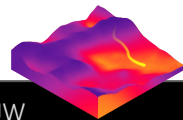




Tabela z wynikami

Method	Depth	Params	C10	C10+	C100	C100+	SVHN
Network in Network [22]	-	-	10.41	8.81	35.68	-	2.35
All-CNN [32]	-	-	9.08	7.25	-	33.71	-
Deeply Supervised Net [20]	-	-	9.69	7.97	-	34.57	1.92
Highway Network [34]	-	-	-	7.72	-	32.39	-
FractalNet [17]	21	38.6M	10.18	5.22	35.34	23.30	2.01
with Dropout/Drop-path	21	38.6M	7.33	4.60	28.20	23.73	1.87
ResNet [11]	110	1.7M	-	6.61	-	-	-
ResNet (reported by [13])	110	1.7M	13.63	6.41	44.74	27.22	2.01
ResNet with Stochastic Depth [13]	110	1.7M	11.66	5.23	37.80	24.58	1.75
	1202	10.2M	-	4.91	-	-	-
Wide ResNet [42]	16	11.0M	-	4.81	-	22.07	-
	28	36.5M	-	4.17	-	20.50	-
	16	2.7M	-	-	-	-	1.64
ResNet (pre-activation) [12]	164	1.7M	11.26*	5.46	35.58*	24.33	-
	1001	10.2M	10.56*	4.62	33.47*	22.71	-
DenseNet ($k = 12$)	40	1.0M	7.00	5.24	27.55	24.42	1.79
DenseNet ($k = 12$)	100	7.0M	5.77	4.10	23.79	20.20	1.67
DenseNet ($k = 24$)	100	27.2M	5.83	3.74	23.42	19.25	1.59
DenseNet-BC ($k = 12$)	100	0.8M	5.92	4.51	24.15	22.27	1.76
DenseNet-BC ($k = 24$)	250	15.3M	5.19	3.62	19.64	17.60	1.74
DenseNet-BC ($k = 40$)	190	25.6M	-	3.46	-	17.18	-

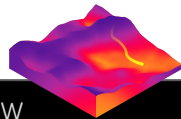




Tabela z wynikami

Method	Depth	Params	C10	C10+	C100	C100+	SVHN
Network in Network [22]	-	-	10.41	8.81	35.68	-	2.35
All-CNN [32]	-	-	9.08	7.25	-	33.71	-
Deeply Supervised Net [20]	-	-	9.69	7.97	-	34.57	1.92
Highway Network [34]	-	-	-	7.72	-	32.39	-
FractalNet [17]	21	38.6M	10.18	5.22	35.34	23.30	2.01
with Dropout/Drop-path	21	38.6M	7.33	4.60	28.20	23.73	1.87
ResNet [11]	110	1.7M	-	6.61	-	-	-
ResNet (reported by [13])	110	1.7M	13.63	6.41	44.74	27.22	2.01
ResNet with Stochastic Depth [13]	110	1.7M	11.66	5.23	37.80	24.58	1.75
	1202	10.2M	-	4.91	-	-	-
Wide ResNet [42]	16	11.0M	-	4.81	-	22.07	-
	28	36.5M	-	4.17	-	20.50	-
	16	2.7M	-	-	-	-	1.64
ResNet (pre-activation) [12]	164	1.7M	11.26*	5.46	35.58*	24.33	-
	1001	10.2M	10.56*	4.62	33.47*	22.71	-
DenseNet ($k = 12$)	40	1.0M	7.00	5.24	27.55	24.42	1.79
DenseNet ($k = 12$)	100	7.0M	5.77	4.10	23.79	20.20	1.67
DenseNet ($k = 24$)	100	27.2M	5.83	3.74	23.42	19.25	1.59
DenseNet-BC ($k = 12$)	100	0.8M	5.92	4.51	24.15	22.27	1.76
DenseNet-BC ($k = 24$)	250	15.3M	5.19	3.62	19.64	17.60	1.74
DenseNet-BC ($k = 40$)	190	25.6M	-	3.46	-	17.18	-

- co znaczą te liczby? np. tutaj:
 - Error rate klasyfikacji
- opis skrótów, np. tutaj:
 - C10 -- CIFAR-10
 - C10+ -- CIFAR-10 z augmentacją
 - ...
- podświetlenie wyników papera
- można wybrać najpierw tylko jeden dataset jak tabelka jest za duża

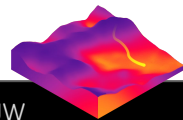




Tabela z wynikami

Method	Depth	Params	C10	C10+
Network in Network [22]	-	-	10.41	8.81
All-CNN [32]	-	-	9.08	7.25
Deeply Supervised Net [20]	-	-	9.69	7.97
Highway Network [34]	-	-	-	7.72
FractalNet [17]	21	38.6M	10.18	5.22
with Dropout/Drop-path	21	38.6M	7.33	4.60
ResNet [11]	110	1.7M	-	6.61
ResNet (reported by [13])	110	1.7M	13.63	6.41
ResNet with Stochastic Depth [13]	110	1.7M	11.66	5.23
	1202	10.2M	-	4.91
Wide ResNet [42]	16	11.0M	-	4.81
	28	36.5M	-	4.17
	16	2.7M	-	-
ResNet (pre-activation) [12]	164	1.7M	11.26*	5.46
	1001	10.2M	10.56*	4.62
DenseNet ($k = 12$)	40	1.0M	7.00	5.24
DenseNet ($k = 12$)	100	7.0M	5.77	4.10
DenseNet ($k = 24$)	100	27.2M	5.83	3.74
DenseNet-BC ($k = 12$)	100	0.8M	5.92	4.51
DenseNet-BC ($k = 24$)	250	15.3M	5.19	3.62
DenseNet-BC ($k = 40$)	190	25.6M	-	3.46

- co znaczą te liczby? np. tutaj:
 - Error rate klasyfikacji
- opis skrótów, np. tutaj:
 - C10 -- CIFAR-10
 - C10+ -- CIFAR-10 z augmentacją
 - ...
- podświetlenie wyników papera
- można wybrać najpierw tylko jeden dataset jak tabelka jest za duża
- wnioski obok tabelki!

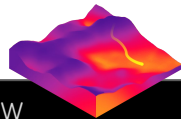
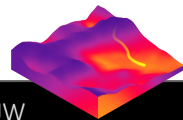




Tabela z wynikami

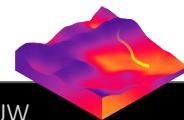
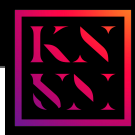
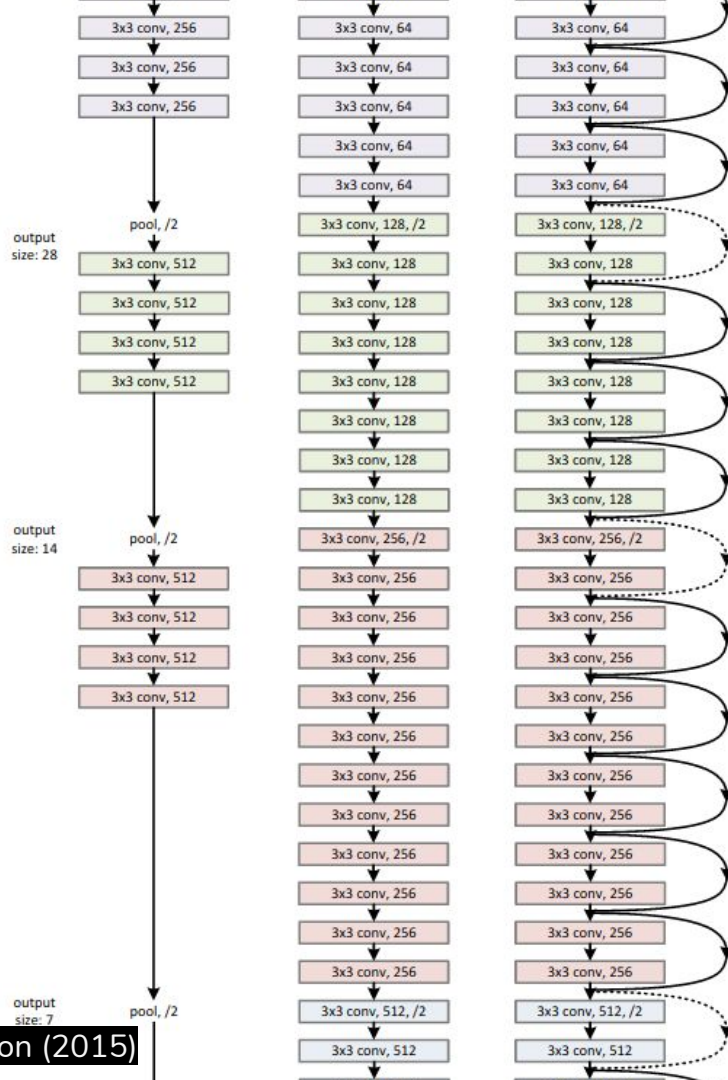
Method	Depth	Params	C10	C10+
Network in Network [22]	-	-	10.41	8.81
All-CNN [32]	-	-	9.08	7.25
Deeply Supervised Net [20]	-	-	9.69	7.97
Highway Network [34]	-	-	-	7.72
FractalNet [17]	21	38.6M	10.18	5.22
with Dropout/Drop-path	21	38.6M	7.33	4.60
ResNet [11]	110	1.7M	-	6.61
ResNet (reported by [13])	110	1.7M	13.63	6.41
ResNet with Stochastic Depth [13]	110	1.7M	11.66	5.23
	1202	10.2M	-	4.91
Wide ResNet [42]	16	11.0M	-	4.81
	28	36.5M	-	4.17
	16	2.7M	-	-
ResNet (pre-activation) [12]	164	1.7M	11.26*	5.46
	1001	10.2M	10.56*	4.62
DenseNet ($k = 12$)	40	1.0M	7.00	5.24
DenseNet ($k = 12$)	100	7.0M	5.77	4.10
DenseNet ($k = 24$)	100	27.2M	5.83	3.74
DenseNet-BC ($k = 12$)	100	0.8M	5.92	4.51
DenseNet-BC ($k = 24$)	250	15.3M	5.19	3.62
DenseNet-BC ($k = 40$)	190	25.6M	-	3.46

- lepsze wyniki na datasetach o różnych trudnościach
- DenseNet osiąga **error rate nawet 40% mniejszy** niż inne rozwiązania na CIFAR-10
- Przy tej samej liczbie parametrów, DenseNet osiąga lepsze wyniki



Skup się na rzeczach istotnych

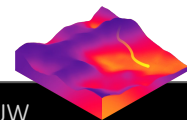
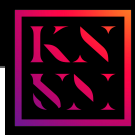
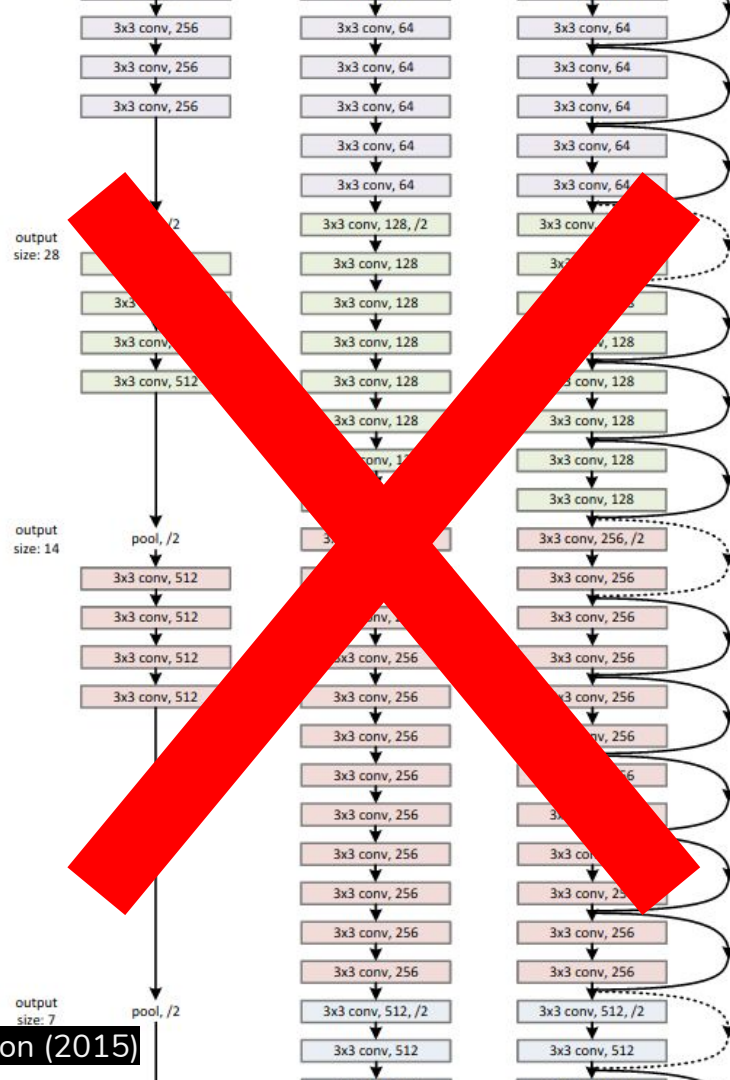
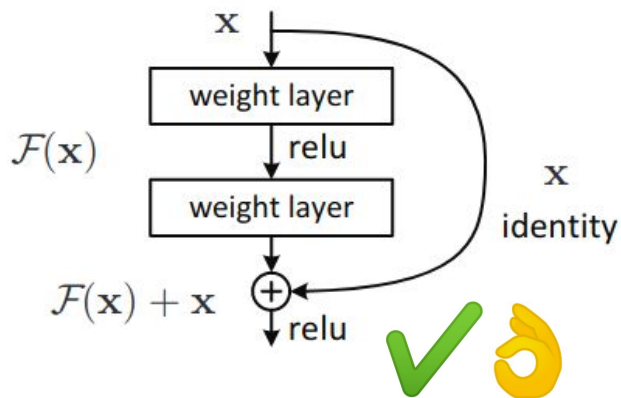
- Pomiń nieistotne szczegóły!
- ...ale przygotuj się na ewentualne pytania



JW

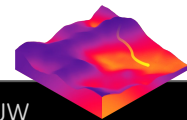
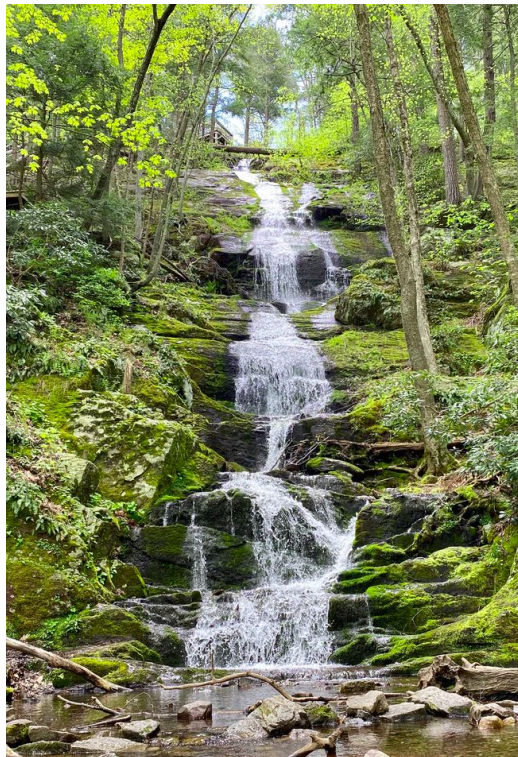
Skup się na rzeczach istotnych

- Pomiń nieistotne szczegóły!
- ...ale przygotuj się na ewentualne pytania



Podawaj źródła!

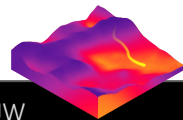
- Źródła grafik / tabel / danych
- w podpisie na slajdzie
- albo na końcu w bibliografii
- albo jak używacie tylko jednego papera to jakoś na początku
- Oczywiście do zdjęć także warto dawać źródła!





Podsumowując

- skup się na przekazaniu głównej idei, a nie szczegółów
- słuchacze mają ograniczony attention span
- słuchaj tych rad tylko jeśli mają sens w Twoim przypadku





Dzięki!

