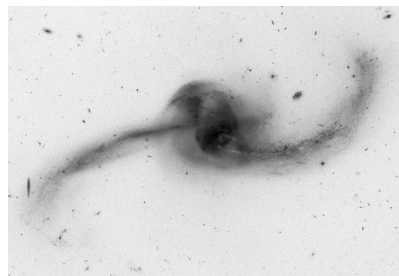


Poszukiwanie zderzających się galaktyk *William J. PEARSON**

* Adiunkt, Narodowe Centrum Badań Jądrowych



Rys. 1. Zdjęcie galaktyki NGC 2623 znajdującej się w odległości około 250 milionów lat świetlnych w gwiazdozbiore Raka. Ten kosmiczny węzeł powstał ze splatania się dwóch galaktyk spiralnych. Źródło: NASA/ESA Hubble Space Telescope

Współczesne duże przeglądy nieba mogą dostarczyć katalog danych zawierający nawet 1 miliard obiektów. Powiedzmy, że nasz astronom poszukując zderzających się galaktyk, poświęca 2 sekundy na każde zdjęcie obiektu z takiego katalogu (co i tak jest optymistycznym założeniem). Pracuje bez przerwy, bez snu i przerw na jedzenie. Obejrzenie wszystkich zdjęć zajęłoby mu 555 555,5 godziny, czyli około 63 lata.

Więcej o współczynniku Giniego można przeczytać w artykule Martyny Kobus *Jak się mierzy nierówności społeczno-ekonomiczne?* (Δ_{18}^{05}).

Zderzenia galaktyk to jedne z najbardziej złożonych zjawisk we Wszechświecie. Nie tylko zachwycają swoim pięknem, tworząc niezwykle struktury (jak dla przykładu ta pokazana na rys. 1), ale również stanowią podstawę współczesnych teorii powstawania i ewolucji struktur we Wszechświecie. Znajdujące się odpowiednio blisko galaktyki przyciągają się wzajemnie, wykonując cudownie skomplikowany, trwający miliardy lat taniec, zakończony długo wyczekiwany uściskiem w spiralnych ramionach partnera. Podczas tych kosmicznych płaśów tempo, w jakim obie galaktyki tworzą nowe gwiazdy, gwałtownie wzrasta i każdego roku rozbłyskują one światłem tysięcy nowych gwiazd. Ten kosmiczny balet sprawia, że gaz i pył opadają do wnętrza galaktyk, karmiąc znajdującą się tam supermasywną czarną dziurę, która po jakimś czasie zaczyna uwalniać niesamowite ilości energii, tworząc obiekt znany jako aktywne jądro galaktyki. Mimo tego wszystkiego zidentyfikowanie galaktyk będących w trakcie zderzenia jest wciąż... zaskakująco trudnym zadaniem.

Klasyfikacja ludzkim okiem

Najbardziej intuicyjną metodą poszukiwania zderzających się galaktyk jest szukanie ich za pomocą „szkiełka i oka”. Postępujące zderzenie galaktyk powoduje ogromne zmiany w strukturze uczestniczących w nim galaktyk, można je więc całkiem łatwo znaleźć na zdjęciach – po prostu kształt takich zlewających się galaktyk bardzo różni się od innych obiektów. Ludzkie oko i mózg doskonale wychwytyują te różnice (przy założeniu, że mamy wystarczająco dobre zdjęcie). Dlatego zidentyfikowane w ten sposób pary zderzających się galaktyk stanowią podstawę, swego rodzaju wzór, dla innych metod. Oczywiście ten rodzaj klasyfikacji ma sporo wad.

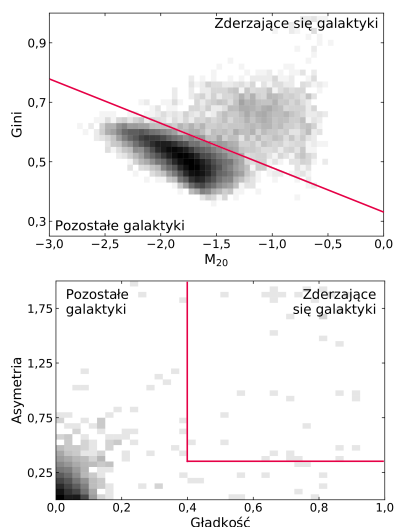
Po pierwsze klasyfikacja wzrokowa jest powolna: wykwalifikowany astronom jest w stanie znaleźć zderzającą się parę galaktyk na zdjęciu w przeciągu kilkudziesięciu sekund. Niby szybko, ale takich galaktyk są przecież miliony! Po drugie katalog zderzających się galaktyk stworzony przez człowieka jest zazwyczaj niemożliwy do idealnego odtworzenia. Klasyfikatorzy nierzadko nie zgadzają się ze sobą, czy dwa obiekty faktycznie są w trakcie zderzenia, zwłaszcza jeżeli mówimy o zdjęciach odległych, których rozdzielczość uniemożliwia dostrzeżenie dokładnych struktur. Co więcej, nawet ta sama osoba może zaklasyfikować obiekt jako zderzające się galaktyki, ale zmienić zdanie następnego dnia lub nawet w przeciągu godziny!

Te wady mogą być nieco złagodzone przez astronomiczne projekty społeczne, takie jak np. Galaktyczne Zoo (*Galaxy Zoo*). W projekcie tym duża liczba wolontariuszy (nie-astronomów) jest proszona o identyfikację tych samych obiektów astronomicznych (np. zderzających się galaktyk). Gdy wszyscy, bądź zdecydowana większość, klasyfikatorzy-amatorzy zgodzą się, że obraz przedstawia zderzające się galaktyki, to możemy mieć niemal pewność, że tak jest naprawdę. Ale nawet twórcy galaktycznego zoo przyznają, że obserwacji, a co za tym idzie zdjęć galaktyk, przybywa w zdecydowanie szybszym tempie niż wolontariuszy. Więc i ta metoda wkrótce przestanie być efektywna.

Klasyfikacja przy użyciu parametrów opisujących kształt galaktyk

Ponieważ klasyfikacja za pomocą wzroku jest powolna i podatna na błędy (a dodatkowo umożliwia identyfikację zdarzających się galaktyk w bardzo specyficznych stadiach zderzenia), astronomowie opracowali metody zautomatyzowane. Jedną z nich uwzględnia kształt (lub inaczej morfologię) galaktyk. Kształt galaktyki może być opisany przy użyciu wielu różnych parametrów, jednak tylko cztery z nich są najczęściej wykorzystywane do klasyfikacji zderzeń. Są to parametry: Gini, M_{20} , asymetrii i gładkości.

O parametrach Gini i M_{20} można by napisać osobne artykuły, tutaj postaram się pokrótce wytłumaczyć, czym są. Współczynnik Giniego, niezbędny do opisu zderzających się galaktyk, ma swoje korzenie w ekonomii! Jest to liczba, która opisuje nierównomierność dowolnego rozkładu. Używana w astronomii



Rys. 2. Przykład klasyfikacji zderzających się galaktyk na podstawie parametrów Gini i M_{20} (na górze) oraz parametrów asymetrii i gładkości (na dole). W każdym przypadku linia wyznacza granice klasyfikacji obiektów jako zderzające się galaktyki. Intensywność każdego „pikselu” odpowiada liczbie galaktyk o określonych przez niego parametrach – im ciemniejszy odcień, tym więcej galaktyk

opisuje rozkład jasności światła emitowanego przez galaktykę rozłożonego na poszczególne piksele obrazu tej galaktyki. Współczynnik Giniego równy 1 opisuje sytuację, w której całe światło galaktyki zawarte jest w jednym pikselu, podczas gdy Gini równy 0 opisuje sytuację, w której światło jest równomiernie rozłożone na wszystkie piksele obrazu. Drugi z parametrów, M_{20} , również opisuje rozkład jasności światła galaktyki. Obliczając ten parametr, mierzy się odległości 20% najjaśniejszych pikseli galaktyki od jej centrum i mnoży się je przez natężenie światła w tych pikselach. Suma uzyskanych wartości daje M_{20} . Parametr ten jest zawsze ujemny, przy czym wartości bliższe zeru wskazują, że światło jest bardziej skoncentrowane w centrum galaktyki.

Kolejne dwa parametry, asymetria i gładkość (jednorodność), są zdecydowanie bardziej intuicyjne. Parametr asymetrii opisuje (rzecz jasna), jak bardzo asymetryczna jest galaktyka. Aby określić ten parametr, stosuje się dość sprytną metodę. Obraca się zdjęcie galaktyki o 180° i odejmuje obróconą wartość od oryginalnego obrazu – piksel po pikselu. Im większa różnica pomiędzy obrazami, tym galaktyka jest bardziej asymetryczna. Z kolei parametr gładkości opisuje, jak jednorodna jest galaktyka. Na obraz nakładany jest filtr wygładzający, a następnie wygładzony obraz jest odejmowany od oryginału. Podobnie jak w przypadku asymetrii, większa różnica wskazuje na mniej jednorodną galaktykę.

Analizując liniowe kombinacje wartości Gini i M_{20} danej galaktyki lub wartości asymetrii i gładkości, astronomowie identyfikują galaktyki będące w czasie zderzenia. Przyjmuje się, że są to obiekty, dla których $Gini > -0,15 \cdot M_{20} + 0,33$ (jak pokazano na rys. 2) lub gładkość $> 0,4$ i asymetria $> 0,35$. Metoda ta jest zdecydowanie szybsza niż klasyfikacja ludzka, a jej wyniki są powtarzalne i niesubiektywne.

Klasyfikacja automatyczna wykonana za pomocą opisanego powyżej zestawu parametrów nie jest oczywiście pozbawiona problemów. Parametry są wrażliwe na jakość i rozdzielczość obrazu: niska rozdzielczość i zaszumiony obraz dają niedokładne wartości, dodatkowo obarczone błędem pomiaru. Wiadomo również, że zakresy wartości parametrów używane do kwalifikacji zderzających się galaktyk często opisują zwykle pojedyncze galaktyki o nieregularnym kształcie.

Poszukiwanie bliskich par galaktyk

Założenie tej metody jest bardzo proste: galaktyki, które są blisko siebie mają sporą szansę się zderzyć. Astronomowie poszukują więc takich bliskich par galaktyk. Jak? Na podstawie ich pozycji na niebie i parametru przesunięcia ku czerwieni (czyli parametru określającego w pewien sposób odległość, pisaliśmy o nim w Δ_{19}^{04}) wyznacza się odległość w przestrzeni trójwymiarowej oraz prędkość względną galaktyk. Ta ostatnia wielkość jest używana do określenia, czy galaktyki ostatecznie zderzą się, czy po prostu przelecą obok siebie. Kryteria identyfikacji bliskich par galaktyk są różne, ale zazwyczaj wybierane są galaktyki, które są oddalone od siebie o co najwyżej 50 kpc ($1,5 \cdot 10^{18}$ km) i mają prędkość względną mniejszą niż 500 km/s.

Dla przykładu średnica Drogi Mlecznej to około 30 kpc, natomiast odległość Drogi Mlecznej od Galaktyki Andromedy – około 770 kpc.

Metoda ta ma jednak wiele wad. Ponieważ klasyfikuje tylko bardzo bliskie pary galaktyk, pomijane są bardzo wczesne etapy łączenia się galaktyk, gdy galaktyki są

od siebie oddalone o więcej niż 50 kpc. Podobnie tracone są późne etapy zderzenia, gdy galaktyki są już tak blisko siebie, że trudno je odseparować na zdjęciach, lub gdy galaktyki zderzyły się, ale struktura nowo powstałej galaktyki nadal jest zaburzona. Pomimo limitu względnej prędkości często błędnie klasyfikowane są galaktyki, które tylko przelatują obok siebie. Ponadto metoda ta wymaga niezwykle precyzyjnych pomiarów przesunięć ku czerwieni, a te są kosztowne (zarówno pod względem czasowym, jak i finansowym) i wymagają wysoce specjalistycznego sprzętu.

Metody uczenia maszynowego

W ostatnich latach popularne stały się metody uczenia maszynowego – obecnie szeroko wykorzystywane również w astronomii. Stosuje się je do klasyfikacji typów galaktyk, poszukiwania aktywnych jąder galaktyk, a od niedawna także do identyfikacji zderzających się galaktyk. Metody te w klasycznej odsłonie wykorzystują obrazy rzeczywistych lub symulowanych galaktyk, tych zderzających się i niezderzających się, do trenowania szeregu algorytmów. Ostatnio jednak poza zdjęciami galaktyk uwzględnia się również dodatkowe dane, takie jak parametry morfologiczne lub przesunięcie ku czerwieni.

Dane treningowe mogą pochodzić z katalogów stworzonych z wykorzystaniem wcześniej opisanych metod: klasyfikacji ludzkiej, klasyfikacji morfologicznej, metody bliskich par lub ich dowolnych kombinacji. Po takim treningu algorytmy są w stanie klasyfikować galaktyki w podobny sposób jak którakolwiek z tych metod, ale zdecydowanie bardziej efektywnie.

Ponadto techniki te są w stanie przezwyciężyć większość niedociągnięć innych metod. Ponieważ uczenie maszynowe jest algorytmiczne, raz wytrenowane narzędzie uczenia maszynowego zawsze dostarczy taki sam wynik dla tych samych danych, eliminując jedną z głównych wad klasyfikacji ludzkiej – brak powtarzalności. Ale to nie wszystko. Metody uczenia maszynowego dobrze radzą sobie z problemami związanymi z jakością obrazu, przy założeniu, że dane treningowe mają taki sam zakres jakości i rozdzielczości obrazu. Co więcej, w metodach uczenia maszynowego nie jest wymagana wysoka precyzja wyznaczenia przesunięć ku czerwieni, przez co łagodzone są koszty obserwacyjne. Aż w końcu, po przeszkoleniu, metody uczenia maszynowego są o rzędy wielkości szybsze niż klasyfikacja ludzka.

Oczywiście klasyfikacje oparte na uczeniu maszynowym nie są pozbawione wad. Nawet najlepsze metody uczenia maszynowego mogą mieć trudności z powodu wysokiego stopnia zanieczyszczenia próbek. Nieprawidłowa klasyfikacja może sięgać nawet 80% zidentyfikowanych galaktyk. Metody uczenia maszynowego są również z natury swej wrażliwe na jakość próbki treningowej.

Jeśli dane treningowe nie zawierają np. bardzo bliskich par, które są pomijane przez metodę par, wówczas wyszkolone narzędzie uczenia maszynowego nie jest w stanie zidentyfikować tego rodzaju zderzeń, gdyż po prostu algorytm nie wie o ich istnieniu. Oznacza to, że metody uczenia maszynowego w większości przypadków nie są w stanie dokonać zaskakujących odkryć, do których zdolny jest ludzki klasyfikator.

Tak więc istnieje wiele różnych sposobów na identyfikację zderzających się galaktyk. Każda z metod ma swoje wady i zalety. W najbliższej przyszłości spodziewamy się wzrostu wykorzystania technik uczenia maszynowego – ze względu na ich wszechstronność, szybkość i potencjalną niezawodność. Metody te wymagają jednak danych do trenowania, a to wymusza użycie i rozwój także innych metod klasyfikujących. Jednak niezależnie od tego, jakie metody będą stosowane w przyszłości, nawet te jeszcze niewymyślone, wciąż istnieją galaktyki, które wymagają identyfikacji, oraz fascynujące odkrycia, których można dokonać dzięki analizie procesów, jakim podlegają zderzające się galaktyki. Musimy je tylko znaleźć.

Tłumaczenie: Anna DURKALEC



Zadania

Przygotował Dominik BUREK

M 1711. Udowodnij, że dowolne dwa dzielniki $d' > d$ liczby n spełniają nierówność

$$d' > d + \frac{d^2}{n}.$$

Rozwiązanie na str. 4

M 1712. $2n$ promieni podzieliło okrąg na $2n$ równych części: n niebieskich i n czerwonych, pokolorowanych w losowej kolejności. W niebieskich sektorach, zaczynając od pewnego, liczby od 1 do n są zapisywane w kierunku przeciwnym do ruchu wskazówek zegara. W czerwonych sektorach, zaczynając od pewnego, zapisane są te same liczby, ale w kierunku zgodnym z ruchem wskazówek zegara. Udowodnij, że istnieje półkole zawierające wszystkie liczby od 1 do n .

Rozwiązanie na str. 4

M 1713. Dany jest trapez $ABCD$ ($AB \parallel CD$) wpisany w okrąg Ω . Punkty M i N są środkami ramion odpowiednio AD i BC . Punkt P jest dowolnym punktem leżącym na krótszym łuku AD okręgu Ω . Udowodnij, że

$$PN + PM \geq BD.$$

Rozwiązanie na str. 9

Przygotował Andrzej MAJHOFER

F 1049. W wyniku przemiany adiabaticznej objętość V jednoatomowego gazu wzrosła q razy. Jak, w wyniku tej przemiany, zmieniła się częstość zderzeń atomów gazu z jednostkową powierzchnią ścianek naczynia, w którym ten gaz się znajduje?

Rozwiązanie na str. 15

F 1050. W przemianie β^+ jądro atomowe krzemu ^{27}Si ($Z_1 = 14$) rozpada się na lustrzane doń jądro ^{27}Al ($Z_2 = 13$) i pozyton e^+ oraz neutrino. Maksymalna energia emitowanego pozytonu wynosi $\Delta E = 3,48$ MeV (wówczas energia neutrino jest zaniedbywalnie mała). Przy założeniu, że ładunek jest rozłożony równomiernie wewnątrz sferycznego jądra o promieniu R , oszacuj, ile wynosi ten promień.

Wskazówka: energia elektrostatyczna jednorodnie naładowanej kuli o ładunku Q i promieniu R wynosi

$$E = \frac{3}{5} \cdot \frac{Q^2}{4\pi\epsilon_0 R},$$

gdzie $\epsilon_0 = 8,854 \cdot 10^{-12}$ F/m jest przenikalnością elektryczną próżni. Ładunek elementarny $e = 1,602 \cdot 10^{-19}$ C, a promień jądra zależy jedynie od jego liczby masowej A : $R = r_0 A^{1/3}$.

Rozwiązanie na str. 15