

Rachkovskij D.A., Gritsenko V.I. Distributed Representation of Vector Data Based on Random Projections. Kiev: Interservice, 2018. 216 p. (in Ukrainian)

Рачковський Д.А., Гриценко В.І. Розподілене подання векторних даних на основі випадкових проєкцій. Київ: Інтерсервіс, 2018. 216 с.

Монографію присвячено методам і алгоритмам перетворення початкових векторних даних в розподілені векторні подання. Розглянуто методи без навчання, які використовують випадкові проєкції (random projection). Отримані розподілені векторні подання можуть застосовуватися як для швидкого оцінювання деяких мір схожості / відстані між векторами великої розмірності (скалярний добуток, евклідова відстань тощо), так і в інших завданнях, які використовують операції з векторами (пошук за схожістю, машинне навчання тощо). Проаналізовано як класичні методи і алгоритми, так і такі, що здобули популярність нещодавно, включаючи результати власних досліджень.

Для науково-технічних працівників, програмістів, аспірантів, студентів та читачів, які цікавляться новими перспективними напрямками інформатики, нейромережовим розподіленим поданням даних, проблематикою швидкого оцінювання схожості.

Рачковский Д.А., Гриценко В.И. Распределенное представление векторных данных на основе случайных проекций. Киев: Интерсервис, 2018. 216 с. (на украинском языке)

Монография посвящена методам и алгоритмам преобразования исходных векторных данных в распределенные векторные представления. Рассмотрены методы без обучения, использующие случайные проекции (random projection). Полученные распределенные векторные представления могут применяться как для быстрой оценки некоторых мер сходства/расстояния между векторами большой размерности (скалярное произведение, евклидово расстояние и др.), так и в других задачах, использующих операции с векторами (поиск по сходству, машинное обучение и др.). Изложены как классические методы и алгоритмы, так и получившие известность недавно, включая результаты собственных исследований.

Для научно-технических работников, программистов, аспирантов, студентов и читателей, интересующихся новыми перспективными направлениями информатики, нейросетевым распределенным представлением данных, проблематикой быстрой оценки сходства.

Empty page
Please see References below

Передмова

Міри схожості і відстані широко застосовують як в пошуку за схожістю, так і в багатьох задачах аналізу даних, статистики, машинного навчання тощо. Наприклад, кластерний аналіз, класифікація та апроксимація методом найближчого сусіда та інші. Будемо називати мірами схожості також і відстані, які визначають міру несхожості.

Якщо точне обчислення схожості/відстані початкових подань об'єктів вимагає великих обчислювальних витрат, то потрібно застосовувати швидке оцінювання. Швидкі оцінки зазвичай є наближеними. Щоб швидко оцінити схожість, часто використовують перетворення початкових подань даних, векторних і неекторних об'єктів різного типу, у векторні подання. Останні дозволяють ефективно оцінити різні міри схожості вхідних даних. Складність обчислення багатьох мір відстані і схожості між векторами лінійно залежить від їхньої розмірності, тому за невеликої розмірності векторів складність невелика. Прикладами таких ефективних мір є евклідова відстань і скалярний добуток. Для бінарних векторів оцінювання схожості ще більш прискорюється за рахунок обчислювальної простоти операцій з бітами. Для векторних подань також є великий арсенал методів пошуку за схожістю, статистичного розпізнавання образів, класифікації, кластеризації, апроксимації, відбору інформативних ознак тощо.

У книзі розглядаються первинні дані у вигляді векторів здебільшого з дійсними компонентами. Для початкових векторів великої розмірності, тобто з великою кількістю компонентів, обчислення мір їхньої схожості є неприпустимо складним (довгим), тому є актуальним швидке наближене оцінювання за перетвореними векторами. Прискорення досягається за рахунок малої розмірності підсумкових дійсних векторів та/або застосування бінарних векторів, оброблення яких є ефективним.

Основна увага приділяється перетворенням, оснований на випадковій проекції, тобто на множенні початкових векторів на матрицю,

елементи якої є реалізаціями випадкових величин з деякого розподілу. До отриманих векторів можуть застосовуватися нелінійні операції, в цій книзі — це операції квантування окремих компонентів, зокрема, операція бінаризації за допомогою порівняння з порогом. Залежно від конкретного типу випадкової проекції та подальшого перетворення, за отриманими векторами можна оцінити певні міри схожості/відстані вхідних векторів.

Зазначимо, що зміст компонентів початкового вектора зазвичай є очевидним: компонент відповідає деякій відомій ознаці (атрибуту) об'єкта, а величина компонента — це значення ознаки для конкретного об'єкта. У вихідних векторах зміст компонентів вже інший, і його поняттява інтерпретація часто складніша і неочевидна. Так, у разі випадкової проекції обчислюється скалярний добуток початкового вектора і деякого випадкового вектора, тобто лінійна комбінація компонентів початкового вектора з випадковими коефіцієнтами. У подальшому нелінійному перетворенні ця величина, наприклад, порівнюється з граничним значенням, і значення компонента підсумкового вектора встановлюється в залежності від результату порівняння. Такі векторні подання, в яких зміст компонентів неочевидний, але за якими можна оцінити деякі міри схожості/відстані початкових об'єктів, ми відносимо до типу «розподілених подань». Вважається, що розподілені подання можуть мати свої аналоги в нейромережевому поданні даних в мозку.

Книга складається із чотирьох розділів.

У розділі 1 розглянуто основні поняття, які використовуються в галузі перетворення початкового подання об'єктів в векторні подання та швидкого наближеного оцінювання вхідних мір схожості і відстані за отриманими векторами.

Дано визначення метричних і векторних просторів. Наведено широко вживані міри схожості/відстані векторів. Розглянуто джерела отримання векторного подання великої розмірності з первинних даних. Обговорюються підходи до формування з таких векторів перетворених векторних подань, які отримали назву «вкладень» і «скетчів». Для вкладень оцінювання схожості/відстані виконується за мірами схожості/відстані між векторами вкладення, а для скетчів можуть застосовуватися інші типи оцінювання, наприклад, медіанні. Вкладення і скетчі дають змогу швидко, хоча і в більшості випадків наближено, оцінити міри схожості/відстані початкових подань. Також

розглянуто розподілені і локальні векторні подання даних, їхні переваги та недоліки.

У розділі 2 обговорено формування векторних подань з дійсними компонентами, яке використовує випадкові проекції.

Розглянуто формування векторів малої розмірності у разі випадкової проекції гаусовими випадковими матрицями, а також випадковими матрицями з елементами з інших субгаусових розподілів. Це дозволяє з малим спотворенням і достатньою точністю оцінити евклідову відстань і скалярний добуток початкових векторів за евклідовою відстанню і скалярним добутком отриманих дійсних векторів малої розмірності, тобто виконується зниження розмірності. Обговорюється точність оцінювання, як для гіршого випадку, так і в термінах дисперсії, та пришвидшення випадкової проекції. Відзначимо, що збереження величин скалярного добутку дає змогу застосовувати отримані вектори у різних лінійних моделях.

Викладено результати про неможливість зниження розмірності для неевклідових відстаней Мінковського не тільки випадковою проекцією, а й іншими перетвореннями. Ця проблема вирішується шляхом оцінювання не за відстанями між отриманими векторами, а застосуванням інших оцінок. Наприклад, медіанних, тобто оцінювання за скетчами. Розглянуто дійсні векторні подання для апроксимації ядерної схожості, для оцінювання статистичних відстаней і для пошуку за схожістю. Обговорюються переваги і недоліки випадкової проекції, надано порівняння з методами з навчанням.

У розділі 3 викладено підходи до формування бінарних і цілочисельних векторів на основі випадкових проекцій.

Розглянуто бінаризовані вкладення для оцінювання кута між початковими векторами, які отримують за покомпонентною бінаризацією (порівнянням з нулем) дійсних компонентів вектора-результату випадкової проекції, а також прискорення формування результату випадкової проекції для цього випадку. Описано універсальний підхід до оцінювання схожості за допомогою векторів з дискретними компонентами, отриманими за локально-чутливим хешуванням (locality-sensitive hashing). Оцінювання може проводитися за допомогою скалярного добутку векторів, що також дає змогу використовувати отримані вектори в лінійних моделях.

Розглянуто бінарні вектори для апроксимації ядерних схожостей, а також цілочисельні скетчі, які сформовано квантуванням

випадкових проєкцій зі стійких розподілів, для оцінювання деяких відстаней Мінковського. Описано скетчі для оцінювання мір схожості початкових бінарних векторів, а також скетчі, отримані цілочисельним квантуванням, скетчі для пошуку за схожістю тощо. Обговорюється ефективність застосування бінарних векторів, порушено питання порівняння з методами, які використовують адаптацію до даних.

Розділ 4 присвячено формуванню бінарних розріджених векторів (тобто з малою часткою одиничних компонентів), так званих кодвекторів, за допомогою випадкової проєкції та бінаризації з додатною величиною порога. Скалярний добуток кодвекторів і інші оцінки на його основі дають можливість оцінити схожість початкових дійсних векторів. Такі кодвекторні подання застосовують в асоціативно-проективних нейронних мережах, розподіленій асоціативній пам'яті матричного типу, обчислювальній інфраструктурі пошукових систем тощо. Наведено варіанти нейромережевої та комп'ютерної реалізацій. Також викладено загальний підхід до формування кодвекторів і оцінювання за ними схожості.

У монографії систематизовано основні напрями досліджень в галузі прискорення оцінювання мір схожості векторних подань з використанням випадкових проєкцій, включаючи власні розробки. Крім базових підходів і методів, розглянуто нові теоретичні розробки, методи і алгоритми. Багато методів та алгоритмів описано докладно, що дає змогу підготовленим читачам використати їх для програмної реалізації. Для деяких алгоритмів наведено порівняльні результати експериментальних досліджень.

Відзначимо, що питання, пов'язані із застосуванням методів машинного навчання, порушено тільки в підрозділах «висновки», оскільки цей напрям досліджень заслуговує окремого розгляду.

Матеріал монографії та бібліографію може бути використано для розроблення навчальних курсів з інформатики.

Автори висловлюють щиро вдячність за плідні обговорення та слушні поради Л.М. Касаткіній та С.В. Сліпченку, О.Д. Гольцеву, Л.М. Козак, В.В. Луковичу, О.Г. Ревуновій та А.А. Фролову, І.С. Місуну та А.М. Соколову, а також рецензентам.

Післямова

У монографії розглянуто деякі методи і алгоритми для швидкого оцінювання різних мір схожості подань об'єктів. Таке оцінювання виконувалася, головним чином, за векторами, компоненти яких є дійсними або бінарними числами.

Початковими даними (поданнями об'єктів) в більшості випадків були вектори великої розмірності. Обговорено оцінки різних мір схожості вхідних векторів, таких як скалярний добуток, косинус кута, ядерні схожості тощо, і різних мір відстані, таких як евклідова, манхеттенська, статистичні, кут та інші міри.

Викладені методи і алгоритми в основному не використовують адаптації до даних, що перетворюються, тобто навчання. Вони в тому або іншому вигляді застосовують випадкові проекції — множення векторів на випадкову матрицю, тобто є рандомізованими. Компоненти отриманого вектора можуть піддаватися квантуванню, наприклад, бінаризації, що формує вектори із цілочисельними компонентами, наприклад, з бінарними компонентами 0 або 1. Більшість розглянутих методів формування вихідних векторів не залежить від інших векторів, які перетворюються, тобто є «забудькуватими».

Так, високу точність оцінювання евклідових відстаней між вхідними дійсними векторами великої розмірності дає евклідова відстань між дійсними векторами малої розмірності, які отримано гаусовою випадковою проекцією. В цьому випадку швидка оцінка схожості досягається шляхом зменшення розмірності розглянутих векторів. Якщо оцінку вхідної міри відстані вдається отримати за вихідними векторами з бінарними компонентами, то прискорення і економії пам'яті на подання та зберігання можна досягти навіть без зниження розмірності векторів, а лише за рахунок ефективності подання та оброблення бінарних векторів. Наприклад, за бінаризованими дійсними векторами, отриманими випадковою проекцією, можна легко оцінити кут між вхідними векторами.

Швидке оцінювання схожості є потрібним для багатьох застосувань. Воно прискорює виконання наближеного лінійного пошуку за

схожістю [335], [336], який виконується обчисленням величини міри схожості між об'єктом-запитом і всіма об'єктами бази, та вибором найбільш схожих об'єктів бази. Якщо вдається добре апроксимувати вхідні початкові міри схожості, то лінійний пошук за ними може дати прийнятний наближений результат.

Відзначимо, що прискорення лінійного пошуку за схожістю не забезпечує досягнення часу виконання запитів, сублінійного щодо розміру бази. На досягнення сублінійного пошуку орієнтовано індексні структури — структури даних, що розробляються для прискорення пошуку за схожістю [375], [448], [337], [338], [339], [340]. Однак найчастіше прискорення зникає у разі збільшення розмірності векторів, які використовуються для подань об'єктів. Водночас, лінійний пошук за схожістю завжди прискорюється за умови швидкого оцінювання схожості. У цьому випадку немає витрат пам'яті на зберігання індексних структур і вхідних подань, а зберігання компактних векторів, особливо бінарних, вимагає набагато меншого обсягу пам'яті.

Швидке оцінювання схожості є потрібним і в інших задачах аналізу даних і машинного навчання — наприклад, класифікація або апроксимація методом найближчого сусіда або ядерними методами, кластеризація тощо. Отримані векторні подання об'єктів можуть застосовуватися не тільки для швидкого оцінювання міри схожості та в основаних на відстанях/схожостях алгоритмах. Вони також можуть використовуватися (безпосередньо або за належного перетворення) в розробленому саме для векторів арсеналі алгоритмів, таких як методи статистичного розпізнавання образів, лінійні і нелінійні алгоритми класифікації та апроксимації, індексні структури швидкого пошуку за схожістю, відбір інформативних ознак та інші методи машинного навчання.

Типовим прикладом є лінійні моделі. Якщо скалярний добуток отриманих векторів апроксимує значення деякої ядерної схожості між вхідними об'єктами, то вектори можуть застосовуватися для навчання лінійних моделей розв'язанню задач класифікації, апроксимації тощо, замість застосування нелінійних ядерних методів.

Відзначимо, що ядерні методи навчання з нелінійним ядром (наприклад, класифікація за допомогою ядерного методу опорних векторів SVM) забезпечують високу точність розв'язання задач класифікації. Однак вони вимагають застосування матриці ядерних

схожостей між усіма об'єктами бази, тобто квадратичних від розміру бази витрат пам'яті, і для великих баз зберігання такої матриці не піддається реалізації. А повторне обчислення значень ядерної схожості on-demand часто вимагає великих обчислювальних витрат.

У разі використання векторів, які апроксимують ядерну схожість скалярним добутком, для навчання лінійних моделей отримують результати, близькі до результатів відповідного ядерного методу. В цьому випадку для великих навчальних вибірок ефективнішими є методи навчання лінійних моделей, а не обчислювально складні ядерні методи.

Вектори, отримані у разі перетворення вхідних векторів різного типу в дійсні вектори малої розмірності за допомогою випадкової проекції, для спеціальних видів розподілів елементів випадкових матриць дозволяють оцінювати певні міри схожості/відстані вхідних векторів. Так, проекції випадковими матрицями з елементами із субгаусового розподілу дають можливість оцінити евклідову відстань, скалярний добуток і кут за цими мірами для отриманих векторів. А випадкова проекція матрицями з елементами зі стійкого розподілу дозволяє оцінити деякі неевклідові відстані Мінковського, однак для цього потрібно проводити оцінку не по відстаням між отриманими векторами, а, наприклад, використовувати медіанні оцінки. Прискорити випадкову проекцію можна за допомогою розріджених матриць і матриць, елементи яких не є незалежними випадковими величинами.

Алгоритми, спеціалізовані для бінарних або цілочисельних векторних даних, часто перевершують алгоритми для дійсних векторів за швидкістю виконання і використанням пам'яті. Хешування, що зберігає схожість, є універсальним механізмом формування векторів з бінарними або цілочисельними компонентами зі збереженням схожості. Прикладом такого хешування є випадкова проекція матрицями з елементами із субгаусового розподілу з бінаризацією отриманого дійсного вектора порівнянням його компонентів з нульовим порогом. Це дозволяє оцінити кут між вхідними векторами за емпіричною ймовірністю збігу хеш-компонентів отриманих векторів. Іншим прикладом є генерація значень хешів квантуванням з багатьма порогоми після випадкової проекції матрицями з елементами зі стійкого розподілу. Внаслідок чого одержується можливість оцінити деякі типи відстаней

Мінковського між вхідними векторами за емпіричною ймовірністю збігу хешів.

Формовані вектори можна подати такими бінарними векторами, скалярний добуток яких апроксимує міру схожості, яку зберігає хешування. Це уможлиблює застосування таких бінарних векторів для лінеаризації ядерних методів.

Розріджені (з малою кількістю ненульових компонентів) бінарні вектори, так звані кодвектори, є затребуваними завдяки наявності ефективних спеціалізованих для них алгоритмів, а також програмно-апаратних засобів для швидкого оброблення таких векторів. Такі векторні подання ми вважаємо належними до методів розподіленого подання даних, які аналогічні поданням інформації в мозку. У методах випадкової проекції з бінаризацією за пороговою операцією величину розрідженості кодвекторів і характер збереження схожості вхідних векторів запропоновано модифікувати за величиною порога.

Схожість кодвекторів обчислюється на основі скалярного добутку з можливим нормуванням різного типу. Тому схожості кодвекторів є ядерними схожостями та можуть використовуватися в ядерних методах. Кодвектори отримують шляхом нелінійного перетворення вхідних векторів, тому вони дають змогу розв'язувати задачі класифікації, апроксимації тощо, які є нелінійними щодо вхідних векторів.

У ряді випадків вдається точно обчислити значення ядерної схожості вхідних векторів, яке апроксимують кодвектори, та використовувати його в ядерних методах. Однак, як вже зазначалося, безпосереднє використання кодвекторів в лінійних методах може бути обчислювально ефективнішим для великої кількості векторів і давати результати, які можна порівняти з нелінійними ядерними методами.

Відзначимо, що якщо скалярний добуток вхідних числових дійсних векторів точно дорівнює або апроксимує деяку нелінійну ядерну схожість об'єктів, то внаслідок їхнього перетворення в кодвектори можна апроксимувати нові нелінійні ядра. Таким чином, розглянуті в книзі методи формування розподілених векторних подань з використанням випадкових проекцій надають можливість: — швидко оцінити деякі міри схожості вхідних векторів великої розмірності;

— у багатьох випадках отримувати компактніші подання об'єктів за вхідні, що економить пам'ять, якщо не зберігати початкові подання;

— для ряду методів отримувати гарантії на точність оцінки схожості;

— використовувати різні алгоритми, розроблені для векторів (індексні структури для пошуку за схожістю, лінійні і нелінійні моделі для класифікації, апроксимації тощо).

Загальними недоліками цих методів є значні витрати ресурсів пам'яті та/або часу на перетворення вхідних подань та приблизність оцінювання схожості.

Обговорювані в монографії методи формування векторів для швидкого оцінювання схожості в переважній більшості не враховують особливості даних конкретної бази і зовсім не використовують навчання з учителем. Застосування методів машинного навчання в формуванні векторів та пошуку за схожістю є великою самостійною галуззю. Розроблено методи перетворення подань та/або мір схожості/відстані, які використовують навчання та належать до напрямів:

— навчання метрики (metric learning) [221], [40], тобто навчання параметрів мір схожості, що поліпшують якість пошуку з точки зору людини для конкретних баз;

— лінійне та нелінійне зниження розмірності [453], [414], [89];

— формування компактних бінарних векторів з навчанням на основі подань і на основі відстаней [424], [423], [425].

Загальним недоліком методів з навчанням є висока обчислювальна складність. Для деяких алгоритмів нетривіальним є формування векторних подань нових об'єктів, до яких не застосовано навчання. Крім того, методи зниження розмірності з адаптацією до даних не завжди розв'язують задачу збереження вхідних схожостей (наприклад такі, що призначено для візуалізації даних).

На закінчення відзначимо, що є методи і алгоритми оцінювання схожості не тільки векторних подань, а й інших типів вхідних подань об'єктів — рядки, графи, об'єкти загальних метричних просторів тощо. Крім того, для оцінювання використовуються не тільки методи формування вихідних векторів на основі проекцій, але й інші підходи. Ці теми заслуговують окремого розгляду.

References

1. Abdullah A., McGregor A., Kumar R., Vassilvitskii S., Venkatasubramanian S. Sketching, embedding, and dimensionality reduction in information spaces. *JMLR: W&CP*. 2016. Vol. 41. P. 948–956.
2. Abraham I., Bartal Y., Neiman O. Advances in metric embedding theory. *Advances in Mathematics*. 2011. Vol. 228, N 6. P. 3026–3126.
3. Achlioptas D. Database-friendly random projections: Johnson-Lindenstrauss with binary coins. *Journal of Computer and System Sciences*. 2003. Vol. 66, N 4. P. 671–687.
4. Aggarwal C.C. An introduction to frequent pattern mining. In: *Frequent Pattern Mining*. Edited by C.C. Aggarwal, J. Han. Basel: Springer, 2014. P. 1–18.
5. Ahmad S., Hawkins J. How do neurons operate on sparse distributed representations? A mathematical theory of sparsity, neurons and active dendrites. *arXiv:1601.00720*. 13 May 2016.
6. Ahn K. J., Guha S., McGregor A. Graph Sketches: Sparsification, spanners, and subgraphs *Proc. PODS'12*. 2012. P. 5–14.
7. Ailon N., Chazelle B. The Fast Johnson–Lindenstrauss Transform and approximate nearest neighbors. *SIAM J. Comput.* 2009. Vol. 39, N 1. P. 302–322.
8. Ailon N., Liberty E. Fast dimension reduction using rademacher series on dual BCH codes. *Discrete and Computational Geometry*. 2009. Vol. 42, N 4. P. 615–630.
9. Ailon N., Liberty E. An almost optimal unrestricted fast Johnson-Lindenstrauss transform. *ACM Transactions on Algorithms*. 2013. Vol. 9, N 3. P. 21.1–21.12.
10. Ailon N., Rauhut H. Fast and RIP-optimal transforms. *Discrete and Computational Geometry*. 2014. Vol. 52, N 4. P. 780–798.
11. Alahi A., Ortiz R., Vanderghyest P. Freak: Fast retina keypoint. *Proc. CVPR'12*. 2012. P. 510–517.
12. Alaoui A. E., Mahoney M. W. Fast randomized kernel methods with statistical guarantees. *Proc. NIPS'15*. 2015. P. 775–783.
13. Alemdar H., Leroy V., Prost-Boucle A., Petrot F. Ternary neural networks for resource-efficient AI applications. *Proc. IJCNN'17*. 2017. P. 2547–2554.
14. Allen-Zhu Z., Gelashvili R., Micali S., Shavit N. Sparse sign-consistent Johnson–Lindenstrauss matrices: Compression with neuroscience-based constraints. *Proceedings of the National Academy of Sciences (PNAS)*. 2014. Vol. 111. P. 16872–16876.
15. Allen-Zhu Z., Gelashvili R., Razenshteyn I. Restricted isometry property for general p-norms. *Proc. SoCG'15*. 2015. P. 451–460.
16. Amosov N.M. *Modelling of thinking and the mind*. New York: Spartan Books, 1967. 192 p.

17. Amosov N.M., Kasatkin A.M., Kasatkina L.M., Kussul E.M., Talaev S.A., Fomenko V. D. Intelligent behaviour systems based on semantic networks. *Kybernetes*. 1973. Vol. 2, N 4. P. 211–216.
18. Amosov N.M., Kussul E. M., Fomenko V. D. Transport robot with a neural network control system. *Advance papers of the Fourth Intern. Joint Conference on Artificial intelligence*. 1975. Vol. 9. P. 1–10.
19. Andoni A. Nearest Neighbor Search: the Old, the New, and the Impossible. PhD thesis. Massachusetts Institute of Technology. 2009.
20. Andoni A., Charikar M., Neiman O., Nguyen H.L. Near linear lower bounds for dimension reduction in L1. *Proc. FOCS'11*. 2011. P. 315–323.
21. Andoni A., Datar M., Immorlica N., Indyk P., Mirrokni V. Locality-Sensitive Hashing using stable distributions. In: *Nearest-Neighbor Methods in Learning and Vision: Theory and Practice*. Edited by G. Shakhnarovich, T. Darrell, P. Indyk. Cambridge, MA: MIT Press, 2006. P. 61–72.
22. Andoni A., Indyk P. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. *Communications of the ACM*. 2008. Vol. 51, N 1. P. 117–122.
23. Andoni A., Indyk P., Laarhoven T., Razenshteyn I., Schmidt L. Practical and optimal LSH for angular distance. *Proc. NIPS'15*. 2015. P. 1225–1233.
24. Andoni A., Krauthgamer R., Razenshteyn I.P. Sketching and embedding are equivalent for norms. *Siam J. Comput.* 2018. Vol. 47, N 3. P. 890–916.
25. Anisimov A.V., Marchenko O.O., Nasirov E.I. Block-diagonal approach to non-negative factorization of sparse linguistic matrices and tensors of extra-large dimension using the latent Dirichlet distribution. *Cybernetics and Systems Analysis*. 2018. Vol. 54, N 6. P. 853–859.
26. Anisimov A., Marchenko O., Taranukha V., Vozniuk T. Semantic and syntactic model of natural language based on non-negative matrix and tensor factorization. *Proc. NLP'14*. 2014. P. 177–184.
27. Arora S., Lee J. R., Naor A. Euclidean distortion and the sparsest cut. *Proc. STOC'05*. 2005. P. 553–562.
28. Arriaga R.I., Vempala S. An algorithmic theory of learning: Robust concepts and random projection. *Machine Learning*. 2006. Vol. 63, N 2. P. 161–182.
29. Babenko A., Slesarev A., Chigorin A., Lempitsky V. Neural codes for image retrieval. *Proc. ECCV'14*. 2014. P. 584–599.
30. Bach F. Sharp analysis of low-rank kernel matrix approximations. *Proc. COLT'13*. 2013. P. 185–209.
31. Baidyk T., Kussul E., Makeyev O., Vega A., Limited receptive area neural classifier based image recognition in micromechanics and agriculture. *International Journal of Applied Mathematics and Informatics*. 2008. Vol. 2, N 3. P. 96–103.
32. Baydyk T., Kussul E., Hernandez Acosta M. LIRA neural network application for microcomponent measurement. *International Journal of Applied Mathematics and Informatics*. Vol.6, N 4. 2012. P.173–180.
33. Bamberger S., Krahermer F. Optimal fast Johnson-Lindenstrauss embeddings for large data sets. *arXiv:1712.01774*. 6 Dec 2017.

34. Baraniuk R. G., Davenport M., DeVore R. A., Wakin M. A simple proof of the Restricted Isometry Property for random matrices. *Constr. Approx.* 2008. Vol. 28, N 3. P. 253–263.
35. Bartal Y., Gottlieb L.-A. Dimension reduction techniques for ℓ_p ($1 \leq p \leq 2$), with applications. *Proc. SoCG'16*. 2016. P. 16:1–16:15.
36. Bar-Yossef Z., Jayram T. S., Kumar R., Sivakumar D. An information statistics approach to data stream and communication complexity. *J. Comput. Syst. Sci.* 2004. Vol. 68, N 4. P.702–732.
37. Batu T., Ergun F., Sahinalp C. Oblivious string embeddings and edit distance approximations. *Proc. SODA'06*. 2006. P. 792–801.
38. Becker A., Ducas L., Gama N., Laarhoven T. New directions in nearest neighbor searching with applications to lattice sieving. *Proc. SODA'16*. 2016. P. 10–24.
39. Belazzougui D., D. Zhang Q. Edit distance: Sketching, streaming and document exchange. *Proc. FOCS'16*. 2016. P. 51–60.
40. Bellet A., Habrard A., Sebban M. A Survey on metric learning for feature vectors and structured data. *arXiv:1306.6709*. 12 Feb 2014.
41. Bengio Y., Paiement J.-F., Vincent P., Delalleau O., Le Roux N., Ouimet M. Out-of-sample extensions for LLE, Isomap, MDS, Eigenmaps, and Spectral Clustering. *Proc. NIPS'03*. 2003. P. 177–184.
42. Bentkus V. A Lyapunov type bound in \mathbb{R}^d . *Theory of Probability & Its Applications*. 2005. Vol. 49. P. 311–323.
43. Berend G. Sparse coding of neural word embeddings for multilingual sequence labeling. *Transactions of the Association for Computational Linguistics*. 2017. Vol. 5. P. 247–261.
44. Berinde R., Gilbert A. C., Indyk P., Karloff H., Strauss M. J. Combining geometry and combinatorics: A unified approach to sparse signal recovery. *Proc. AAC on CCC'08*. 2008. P. 798–805.
45. Berry A.C. The accuracy of the Gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*. 1941. Vol. 49, N 1. P. 122–136.
46. Bhattacharya R.N., Holmes S. An exposition of Gotze's estimation of the rate of convergence in the multivariate central limit theorem. *arXiv:1003.4254*. 22 Mar 2010.
47. Bhattacharya A., Kar P., Pal M. On low distortion embeddings of statistical distance measures into low dimensional spaces. *Proc. DEXA'09*. 2009. P. 164–172.
48. Bodyanskiy Y., Boiko O., Zaychenko Y., Hamidov G. Evolving hybrid GMDH-Neuro-Fuzzy network and its applications. *Proc. SAIC'18*. 2018. P. 1–6.
49. Borak S., Hardle W., Weron R. Stable distributions. SFB 649 Discussion Paper 2005-008. Humboldt Universitat zu Berlin, Germany, 2005. P.1–28.
50. Bottou L. Large-scale machine learning with stochastic gradient descent. *Proc. COMPSTAT'10*. 2010. P. 177–187.

51. Boufounos P.T., Mansour H., Rane S., Vetro A. Dimensionality reduction of visual features for efficient retrieval and classification. *APSIPA Trans. on Signal and Information Processing*. 2016. Vol. 5. P. e14.
52. Boufounos P.T., Rane S., Mansour H. Representation and coding of signal geometry. *Information and Inference: A Journal of the IMA*. 2017. Vol. 6, N 4. P. 349–388.
53. Broder A.Z. On the resemblance and containment of documents. *Proc. SEQUENCES'97*. 1997. P. 21–29.
54. Broder A. Z., Charikar M., Frieze A. M., Mitzenmacher M. Min-wise independent permutations. *J. Comput. System Sci.* 1998. Vol. 60. P. 327–336.
55. Broder A.Z, Glassman S.C., Manasse M.S., Zweig G. Syntactic clustering of the web. *Computer Networks and ISDN Systems*. 1997. Vol. 29, N 8-13. P. 1157–1166.
56. Bourgain J. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.* 1985. Vol. 52. P. 46–52.
57. Bourgain J., Dirksen S., Nelson J. Toward a unified theory of sparse dimensionality reduction in Euclidean space. *Geometric and Functional Analysis*. 2015. Vol. 25, N 4. P. 1009–1088.
58. Braverman V, Ostrovsky R, Rabani Y. Rademacher chaos, random Eulerian graphs and the sparse JohnsonLindenstrauss transform. *arXiv:1011.2590*. 11 Nov 2010.
59. Brinkman B., Charikar M. On the impossibility of dimension reduction in l_1 . *Journal of the ACM*. 2005. Vol. 52, N 5. P. 766–788.
60. Buldygin V., Moskvichova K. The sub-gaussian norm of a binary random variable. *Theory of Probability and Mathematical Statistics*. 2013. Vol. 86. P. 33–49.
61. Burges C.J. C. Dimension reduction: a guided tour. *Foundations and Trends in Machine Learning*. 2009. Vol. 2, N 4. P. 275–365.
62. Calonder M., Lepetit V., Strecha C., Fua P. Brief: Binary robust independent elementary features. *Proc. ECCV'10*. 2010. P. 778–792.
63. Chambers J. M., Mallows C. L., Stuck B. W. A method for simulating stable random variables. *Journal of the American Statistical Association*. 1976. Vol. 71, N 354. P. 340–344.
64. Chandrasekaran V., Recht B., Parrilo P.A. Willsky A.S. The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*. 2012. Vol. 12, N 6. P. 805–849. см. P. 817.
65. Chandar V. B. Sparse graph codes for compression, sensing, and secrecy. PhD thesis. Massachusetts Institute of Technology, 2010.
66. Charikar M. Similarity estimation techniques from rounding algorithms. *Proc. STOC'02*. 2002. P. 380–388.
67. Charikar M., Sahai A. Dimension reduction in the ℓ_1 norm. *Proc. FOCS'02*. 2002. P. 551–560.
68. Chen D., Phillips J.M. Relative error embeddings of the Gaussian kernel distance. *Proc. ALT'17*. 2017. P. 560–576.

69. Chen L.H.Y., Fang X. Multivariate normal approximation by Stein's method: the concentration inequality approach. *arXiv:1111.4073*. 16 May 2015.
70. Cheng H. Sparse Representation, Modeling and Learning in Visual Recognition. Theory, Algorithms and Applications. London: Springer, 2015. 257 p.
71. Cheng Y., Yu F.X., Feris R.S., Kumar S., Choudhary A., Chang S.-F. An exploration of parameter redundancy in deep networks with circulant projections. *Proc. ICCV'15*. 2015. P. 2857–2865.
72. Chernodub A., Nowicki D. Sampling-based gradient regularization for capturing long-term dependencies in recurrent neural networks. *Proc. ICONIP'16*. 2016. P. 90–97.
73. Choromanska A., Choromanski K., Bojarski M., Jebara T., Kumar S., LeCun Y. Binary embeddings with structured hashed projections. *Proc. ICML'16*. 2016. P. 344–353.
74. Choromanski K., Rowland M., Sarlos T., Sindhvani V., Turner R., Weller A. The Geometry of Random Features. *Proc. AISTATS'18*. 2018. P. 1–9.
75. Choromanski K., Rowland M., Weller A. The unreasonable effectiveness of structured random orthogonal embeddings. *Proc. NIPS'17*. 2017. P. 219–228.
76. Choromanski K., Sindhvani V. Recycling randomness with structure for sublinear time kernel expansions. *Proc. ICML'16*. 2016. P. 2502–2510.
77. Clarkson K.L. Nearest-neighbor searching and metric space dimensions. In: Nearest Neighbor Methods for Learning and Vision: Theory and Practice. Edited by G. Shakhnarovich, T. Darrell, P. Indyk. Cambridge, MA: MIT Press. 2006. P. 15–59.
78. Clarkson K.L., Woodruff D.P. Numerical linear algebra in the streaming model. *Proc. STOC'09*. 2009. P. 205–214.
79. Cohen E. Distance queries from sampled data: Accurate and efficient. *Proc. KDD'14*. 2014. P. 681–690.
80. Cohen M.B., Jayram T.S., Nelson J. Simple Analyses of the Sparse Johnson-Lindenstrauss Transform. *Proc. SOSA'18*. 2018. P. 15:1–5:9.
81. Cohen M.B., Lee Y.T., Musco C., Musco C., Peng R., Sidford A. Uniform sampling for matrix approximation. *Proc. ITCS'15*. 2015. P. 181–190.
82. Cohen M.B., Musco C., Musco C. Input sparsity time low-rank approximation via ridge leverage score sampling. *Proc. SODA'17*. 2017. P. 1758–1777.
83. Conte D., Ramel J.Y., Sidere N., Luqman M.M., Gauzere B., Gibert J., Brun L., Vento M. A comparison of explicit and implicit graph embedding methods for pattern recognition. *Proc. GBRPR'13*. 2013. P. 81–90.
84. Cormode G., Garofalakis M., Haas P.J., Jermaine C. Synopses for massive data: Samples, histograms, wavelets, sketches. *Foundations and Trends in Databases*. 2012. Vol. 4, N 1–3. P. 1–294.
85. Cormode G., Indyk P., Koudas N., Muthukrishnan S. 2002b. Fast mining of massive tabular data via approximate distance computations. *Proc. ICDE'02*. 2002. P. 605–616.
86. Cortes C., Mohri M., Talwalkar A. On the impact of kernel approximation on learning accuracy. *Proc. AISTAT'10*. 2010. P. 113–120.

87. Cristianini N., Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge, UK: Cambridge University Press, 2000. 204 p.
88. Cui Y., Ahmad S., Hawkins J. The HTM spatial pooler—A neocortical algorithm for online sparse distributed coding. *Front. Comput. Neurosci.* 2017. Vol. 11. P.11.1–11.15.
89. Cunningham J., Ghahramani Z. Linear dimensionality reduction: Survey, insights, and generalizations. *Journal of Machine Learning Research.* 2015. Vol. 16. P.2859-2900.
90. Dahlgaard S., Igel C., Thorup M. Nearest neighbor classification using bottom-k sketches. *Proc. BigData'13.* 2013. P. 28–34.
91. Dasgupta A., Kumar R., Sarlos T. A sparse Johnson-Lindenstrauss transform. *Proc. STOC'10.* 2010. P. 341–350.
92. Dasgupta A., Kumar R., Sarlos T. Fast locality sensitive hashing. *Proc. SIGKDD'11.* 2011. P. 1073–1081.
93. Dasgupta S., Stevens C.F., Navlakha S. A neural algorithm for a fundamental computing problem. *Science.* 2017. Vol. 358, N 6364. P. 793–796.
94. Datar M., Immorlica N., Indyk P., Mirrokni V. S. Locality-sensitive hashing scheme based on p-stable distributions. *Proc. SCG'04.* 2004. P. 253–262.
95. Deza M., Deza E. Encyclopedia of Distances. Berlin, Heidelberg: Springer, 2014. 733 p.
96. Dirksen S. Dimensionality reduction with subgaussian matrices: a unified theory. *Foundations of Computational Mathematics.* 2016. Vol. 16, N 5. P. 1367–1396.
97. Dirksen S., Stollenwerk A. Fast binary embeddings with Gaussian circulant matrices: improved bounds. *Discrete & Computational Geometry.* Vol. 60, N3. P. 599–626.
98. Donaldson R., Gupta A., Plan Y., Reimer T. Random mappings designed for commercial search engines. *arXiv:1507.05929.* 21 Jul 2015.
99. Dong W., Charikar M., Li K. Asymmetric distance estimation with sketches for similarity search in high-dimensional spaces. *Proc. SIGIR'08.* 2008. P. 123–130.
100. Duffield N., Lund C., Thorup M. Priority sampling for estimating arbitrary subset sums. *JACM.* 2007. Vol. 54, N 6. P. 32.1–32.39.
101. Durrant R.J., Kaban A. Flip Probabilities for Random Projections of θ -separated Vectors. Technical Report CSR-10-10. School of Computer Science, University of Birmingham. 2010.
102. Engebretsen L., Indyk P., O'Donnell R. Derandomized dimensionality reduction with applications. *Proc. SODA'02.* 2002. P. 705–712.
103. Eshghi K., Kafai M. Support Vector Machines with sparse binary high-dimensional feature vectors. HPE-2016-30. 2016.
104. Eshghi K., Kafai M. The CRO Kernel: Using concomitant rank order hashes for sparse high dimensional randomised feature maps. *Proc. ICDE'16.* 2016. P. 721–730.

105. Esseen C.G. On the Liapunov limit of error in the theory of probability. *Arkiv fur Matematik, Astronomi och Fysik*. 1942. Vol. 28A, N 9. P. 1–19.
106. Esseen C.G. A moment inequality with an application to the central limit theorem. *Skandinavisk Aktuarietidskrift*. 1956. Vol. 39. P. 160–170.
107. Fan R.-E., Chang K.-W., Hsieh C.-J., Wang X.-R., Lin C.-J. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*. 2008. Vol. 9. P. 1871–1874.
108. Faruqui M., Dyer C. Non-distributional Word Vector Representations. *Proc. ACL-IJCNLP'15*. 2015. Vol. 2. P. 464–469.
109. Faruqui M., Tsvetkov Y., Yogatam D., Dyer C., Smith N.A. Sparse Overcomplete Word Vector Representations. *Proc. ACL-IJCNLP'15*. 2015. Vol. 1. P. 1491–1500.
110. Fedoruk J., Schmuland B., Johnson J., Heo G. Dimensionality reduction via the Johnson–Lindenstrauss Lemma: theoretical and empirical bounds on embedding dimension. *The Journal of Supercomputing*. 2018. Vol. 74, N 8. P. 3933–3949.
111. Feng C., Hu Q., Liao S. Random feature mapping with signed circulant matrix projection. *Proc. IJCAI'15*. 2015. P. 3490–3496.
112. Feragen A., Kasenburg N., Petersen J., de Bruijne M., Borgwardt K.M. Scalable kernels for graphs with continuous attributes. *Proc. NIPS'13*. 2013. P. 216–224.
113. Ferdowsi S., Voloshynovskiy S., Kostadinov D., Holotyak T. Fast content identification in highdimensional feature spaces using sparse ternary codes. *Proc. WIFS'16*. 2016. P. 1–6.
114. Fainzilberg L. ECG Averaging Based on Hausdorff Metric. *International Journal of Biomagnetism*. 2003. Vol. 5, N 1. P. 236–237.
115. Fainzilberg L.S. Restoration of a standard sample of cyclic waveforms with the use of the Hausdorff metric in a phase space. *Cybernetics and Systems Analysis*. 2003. Vol. 39, N 3. P. 338–344.
116. Figiel T., Lindenstrauss J., Milman V. D.. The dimension of almost spherical sections of convex bodies. *Acta Math*. 1977. Vol. 139, N 1. P. 53–94.
117. Foggia P., Percannella G., Vento M. Graph matching and learning in pattern recognition in the last 10 years. *Int. J. Pattern Recog. Artif. Intell.* 2014. Vol. 28, N 1. P. 1–40.
118. Forbus K.D., Gentner D., Law K. MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*. 1995. Vol. 19, N 2. P. 141–205.
119. Frady E. P., Kleyko D., Sommer F. T. A theory of sequence indexing and working memory in recurrent neural networks. *Neural Comput.* 2018. Vol. 30, N. 6. P. 1449–1513.
120. Freksen C.B., Kamma L., Larsen K.G. Fully understanding the hashing trick. *Proc. NIPS'18*. 2018.
121. Freksen C.B., Larsen K.G. On using Toeplitz and Circulant matrices for Johnson–Lindenstrauss transforms. *Proc. ISAAC'17*. 2017. P. 32:1–32:12.

122. Frolov A. A., Husek D., Rachkovskij D.A. Time of searching for similar binary vectors in associative memory. *Cybernetics and Systems Analysis*. 2006. Vol. 42, N 5. P. 615–623.
123. Frolov A.A., Rachkovskij D.A., Husek D. On information characteristics of Willshaw-like auto-associative memory. *Neural Network World*. 2002. Vol. 12, N 2. P. 141–158.
124. Gabdullin R.A., Makarenko V.A., Shevtsova I. G. Esseen–Rozovskii type estimates for the rate of convergence in the Lindeberg theorem. *Journal of Mathematical Sciences*. Vol. 234, N 6. P. 847–885.
125. Gallant S.I., Culliton P. Positional binding with distributed representations. *Proc. ICIVC'16*. 2016. P. 108–113.
126. Gallant S. I., Okaywe T.W. Representing objects, relations, and sequences. *Neural Comput.* 2013. Vol. 25, N 8. P. 2038–2078.
127. Gartner T., Lloyd J., Flach P. Kernels and distances for structured data. *Machine Learning*. 2004. Vol. 57, N 3. P.205–232.
128. Gayler R. Multiplicative binding, representation operators, and analogy. In: *Advances in Analogy Research: Integration of Theory and Data from the Cognitive, Computational, and Neural Sciences*. Edited by K. Holyoak, D. Gentner, and B. Kokinov. Sofia, Bulgaria: New Bulgarian University, 1998. P. 405.
129. Gibert J., Valveny E., Bunke H. Embedding of graphs with discrete attributes via label frequencies. *Int. J. Patt. Recogn. Artif. Intell.* 2013. Vol. 27, N 3. P. 1–27.
130. Gionis A., Indyk P., Motwani R. Similarity search in high dimensions via hashing. *Proc. VLDB'99*. 1999. P. 518–529.
131. Gittens A., Mahoney M.W. Revisiting the Nystrom method for improved large-scale machine learning. *Journal of Machine Learning Research*. 2016. Vol. 17, N 1. P. 3977–4041.
132. Goemans M., Williamson D. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *JACM*. 1995. Vol. 42, N 6. P. 1115–1145.
133. Goltsev A.D. Structured neural networks with learning for texture segmentation in images. *Cybernetics and Systems Analysis*. 1991. Vol. 27, N 6. P. 927–936.
134. Goltsev A. Secondary learning in the assembly neural network. *Neurocomputing*. 2004. Vol. 62. P. 405–426.
135. Goltsev A.D. Neural networks with assembly organization. Kiev: Naukova Dumka, 2005. 200 p. (in Russian)
136. Goltsev A., Husek D. Some properties of the assembly neural networks. *Neural Network World*. 2002. Vol. 12, N. 1. P. 15–32.
137. Goltsev A., Gritsenko V. Modular neural networks with Hebbian learning rule. *Neurocomputing*. 2009. Vol. 72, N 10–12. P. 2477–2482.
138. Goltsev A.D., Gritsenko V.I. Algorithm of sequential finding the textural features characterizing homogeneous texture segments for the image segmentation task. *Cybernetics and Computer Engineering*. 2013. N 173. P. 25–34. (in Russian)

139. Goltsev A., Gritsenko V. Modular neural networks with radial neural columnar architecture. *Biologically Inspired Cognitive Architectures*. 2015. Vol. 13. P. 63–74.
140. Goltsev A., Gritsenko V., Husek D. Extraction of homogeneous fine-grained texture segments in visual images. *Neural Network World*. 2017. Vol. 27, N 5. P. 447–477.
141. Goltsev A., Gritsenko V., Kussul E., Baidyk T. Finding the texture features characterizing the most homogeneous texture segment in the image. *Proc. IWANN'15*. 2015. P. 287–300.
142. Goltsev A., Wunsch D.C. Inhibitory connections in the assembly network for texture segmentation. *Neural Networks*. 1998. Vol. 11, N 5. P. 951–962.
143. Gong Y., Sanjiv K., Rowley H.A., Lazebnik S. Learning binary codes for highdimensional data using bilinear projections. *Proc. CVPR'13*. 2013. P. 484–491.
144. Gordon Y. On Milman's inequality and random subspaces which escape through a mesh in \mathbb{R}^n . In: *Geometric Aspects of Functional Analysis*. Edited by J. Lindenstrauss, V.D. Milman. Berlin, Heidelberg: Springer, 1988. P. 84–106.
145. Gorisse D., Cord M., Precioso F. Locality-sensitive hashing for chi2 distance. *IEEE Trans. PAMI*. 2012. Vol. 34, N 2. P. 402–409.
146. Gorodnichy D.O., Reznik A.M. Static and dynamic attractors of autoassociative neural networks. *Proc. ICIAP'97*. 1997. P. 238–245.
147. Gray R.T., Robinson P.A. Stability and structural constraints of random brain networks with excitatory and inhibitory neural populations. *Journal of Computational Neuroscience*. 2009. Vol. 27, N 1. P. 81–101.
148. Grauman K., Fergus R. Learning binary hash codes for large-scale image search. In: *Machine Learning for Computer Vision*. Edited by R. Cipolla, S. Battiato, and G.M. Farinella. Berlin, Heidelberg: Springer, 2013. P. 49–87.
149. Gritsenko V.I., Rachkovskij D.A., Frolov A.A., Gayler R., Kleyko D., Osipov E. Neural distributed autoassociative memories: A survey. *Cybernetics and Computer Engineering*. 2017. N 2 (188). P. 5–35.
150. Gritsenko V.I., Rachkovskij D.A., Goltsev A.D., Lukovych V.V., Misuno I.S., Revunova E.G., Slipchenko S.V., Sokolov A.M., Talayev S.A. Neural distributed representation for intelligent information technologies and modeling of thinking. *Cybernetics and Computer Engineering*. 2013. Vol. 173. P. 7–24. (in Russian)
151. Gritsenko V.I., Rachkovskij D.A., Revunova E.G. Neural distributed representation of vector data in intelligent information technologies. *Cybernetics and Computer Engineering*. 2018. N 4 (194). P. 5–27.
152. Guruswami V., Lee J.R., Razborov A. Almost Euclidean subspaces of L_1^n via expander codes. *Combinatorica*. 2010. Vol. 30, N 1. P. 47–68.
153. Halko N., Martinsson P.-G., Tropp J.A. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*. 2011. Vol. 53, N 2. P. 217–288.
154. Hamid R., Xiao Y., Gittens A., DeCoste D. Compact random feature maps. *Proc. ICML'14*. 2014. P. 19–27.

155. Haviv I., Regev O. The Restricted Isometry Property of subsampled Fourier matrices. *Proc. SODA'16*. 2016. P. 288–297.
156. Har-Peled S., Indyk P., Motwani R. Approximate nearest neighbor: Towards removing the curse of dimensionality. *Theory of Computing*. 2012. Vol. 8, N 1. P. 321–350.
157. Heinly J., Dunn E., Frahm J.-M. Comparative evaluation of binary features. *Proc. ECCV'12*. 2012. P. 759–773.
158. Hinrichs A., Vybiral J. Johnson-Lindenstrauss lemma for circulant matrices. *Random Structures & Algorithms*. 2011. Vol. 39, N 3. P. 391–398.
159. Hoeffding W. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*. 1963. Vol. 58, N 301. P. 13–30.
160. Hofmann T., Scholkopf B., Smola A. Kernel methods in machine learning. *Annals of Statistics*. 2008. Vol. 36, N 3. P. 1171–1220.
161. Hong W., Meng J., Yuan J. Tensorized projection for high-dimensional binary embedding. *Proc. AAAI'18*. 2018. P. 69–76.
162. Hong W., Yuan J. Fried binary embedding: from high-dimensional visual features to high-dimensional binary codes. *IEEE Trans. Image Processing*. 2018. Vol. 27, N 10. P. 4825–4837.
163. Hong W., Yuan J., Das Bhattacharjee S. Fried binary embedding for high-dimensional visual features. *Proc. CVPR'17*. 2017. P. 6221–6229.
164. Hsieh S.-H., Lu C.-S., Pei S.-C. Fast binary embedding via circulant downsampled matrix. *Proc. ICIAP'16*. 2016. P. 1789–1793.
165. Hu Z., Bodyanskiy Y.V., Tyshchenko O.K., Samitova V.O. Possibilistic fuzzy clustering for categorical data arrays based on frequency prototypes and dissimilarity measures. *International Journal of Intelligent Systems and Applications*. 2017. Vol. 9, N 5. P. 55–61.
166. Hu Z., Bodyanskiy Y.V., Tyshchenko O.K., Tkachov V.M. Fuzzy clustering data arrays with omitted observations. *Int. J. Intelligent Systems and Applications*. 2017. Vol. 9, N 6. P. 24–32.
167. Hubara I., Courbariaux M., Soudry D., El-Yaniv R., Bengio Y. Binarized neural networks. *Proc. NIPS'16*. 2016. P. 4107–4115.
168. Iclanzan D., Szilagyı S.M., Szilagyı L. Evolving computationally efficient hashing for similarity search. *Proc. ICONIP'18*. 2018.
169. Indyk P. Algorithmic applications of low-distortion geometric embeddings. *Proc. FOCS'01*. 2001. P. 10–35.
170. Indyk P. Stable distributions, pseudorandom generators, embeddings, and data stream computation. *Journal of the ACM*. 2006. Vol. 53, N 3. P. 307–323.
171. Indyk P. Sketching, Streaming and Sub-linear Space Algorithms. Lecture 3. Massachusetts Institute of Technology. Sep 12, 2007. Scribe: Jelani Nelson <http://stellar.mit.edu/S/course/6/fa07/6.895/materials.html>
172. Indyk P., Matousek J., Sidiropoulos A. Low-distortion embeddings of finite metric spaces. In: *Handbook of Discrete and Computational Geometry*. Boca Raton, FL: CRC Press LLC. 2017. P. 211–231.

173. Indyk P., Motwani R. Approximate nearest neighbors: Towards removing the curse of dimensionality. *Proc. STOC'98*. 1998. P. 604–613.
174. Indyk P., Naor A. Nearest-neighbor-preserving embeddings. *ACM Trans. Algorithms*. 2007. Vol. 3, N 3. P. 31.1–31.12.
175. Indyk P., Razenshteyn I., Wagner T. Practical data-dependent metric compression with provable guarantees. *Proc. NIPS'17*. 2017. P. 2614–2623.
176. Ioffe S. Improved consistent sampling, weighted minhash and L_1 sketching. *Proc. ICDM'10*. 2010. P. 246–255.
177. Jacques L. A quantized Johnson–Lindenstrauss lemma: the finding of Buffon’s needle. *IEEE Trans. Inf. Theory*. 2015. Vol. 61, N 9. P. 5012–5027.
178. Jacques L. Small width, low distortions: quasi-isometric embeddings with quantized sub-gaussian random projections. *IEEE Trans. Inf. Theory*. 2017. Vol. 63, N 9. P. 5477–5495.
179. Jacques L., Cambareri V. Time for dithering: fast and quantized random embeddings via the restricted isometry property. *Information and Inference: A Journal of the IMA*. 2017. Vol. 6, N 4. P. 441–476.
180. Jacques L., Laska J. N., Boufounos P. T., Baraniuk R. G. Robust 1-Bit compressive sensing via binary stable embeddings of sparse vectors. *IEEE Trans. Inf. Theory*. 2013. Vol. 59, N 4. P. 2082–2102.
181. Jacobs D. Multidimensional Scaling: More complete proof and some insights not mentioned in class. 2013.
182. Jagadeesan M. Simple analysis of sparse, sign-consistent JL. *arXiv:1708.02966*. 9 Aug 2017.
183. Jayram T.S., Woodruff D.P. Optimal bounds for Johnson-Lindenstrauss transforms and streaming problems with subconstant error. *ACM Trans. on Algorithms*. 2013. Vol.9, N 3. P. 26.1–26.17.
184. Jegou H., Douze M., Schmid C. Product quantization for nearest neighbor search. *IEEE Trans. PAMI*. 2011. Vol. 33, N 1. P. 117–128.
185. Jegou H., Perronnin F., Douze M., Sanchez J., Perez P., Schmid C. Aggregating local image descriptors into compact codes. *IEEE TPAMI*. 2012. Vol. 34, N 9. 1704–1716.
186. Jiang K., Que Q., Kulis B. Revisiting kernelized locality-sensitive hashing for improved large-scale image retrieval. *Proc. CVPR'15*. 2015. P. 4933–4941.
187. Joachims T., Finley T., Yu C.-N. J. Cutting-plane training of structural SVMs. *Machine Learning*. 2009. Vol. 77, N 1. P. 27–59.
188. Johnson W.B., Lindenstrauss J. Extensions of Lipschitz mapping into Hilbert space. *Contemporary Mathematics*. 1984. Vol. 26. P. 189–206.
189. Johnson W.B., Schechtman G. Embedding l_p into l_1 . *Acta Math*. 1982. Vol. 149, N 1. P. 71–85.
190. Joly A., Buisson O. Random maximum margin hashing. *Proc. CVPR'11*. 2011. P. 873–880.
191. Kaban A. Improved bounds on the dot product under random projection and random sign projection. *Proc. KDD'15*. 2015. P. 487–496.

192. Kafai M., Eshghi K. CROification: Accurate kernel classification with the efficiency of sparse linear SVM. *IEEE Trans. PAMI*. 2019. Vol. 41, N 1. P. 34–48.
193. Kane D.M., Meka R., Nelson J. Almost optimal explicit Johnson-Lindenstrauss families. *Proc. RANDOM'11*. 2011. P. 628–639.
194. Kane D.M., Nelson J. A derandomized sparse JohnsonLindenstrauss transform. *arXiv:1006.3585*. 7 Dec 2010.
195. Kane D.M., Nelson J. Sparser Johnson-Lindenstrauss transforms. *Journal of the ACM*. 2014. Vol. 61, N 1. P. 4:1–4:23.
196. Kane D.M., Nelson J., Woodruff D. P. On the exact space complexity of sketching and streaming small norms. *Proc. SODA'10*. 2010. P. 1161–1178.
197. Kanerva P. Binary spatter-coding of ordered k-tuples. *Proc. ICANN'96*. 1996. P. 869–873.
198. Kanerva P. Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. *Cognitive Computation*. 2009. Vol. 1, N 2. P. 139–159.
199. Kanerva P., Kristoferson J., Holst A. Random indexing of text samples for Latent Semantic Analysis. *Proc. of the Annual Meeting of the Cognitive Science Society*. 2000. Vol. 22. P. 1036.
200. Kapralov M., Potluru V., Woodruff D. How to fake multiply by a Gaussian matrix. *Proc. ICML'16*. 2016. P. 2101–2110.
201. Kar P., Karnick H. Random feature maps for dot product kernels. *Proc. ICAIS'12*. 2012. P. 583–591.
202. Karnin Z., Rabani Y., Shpilka A. Explicit dimension reduction and its applications. *SIAM J. Comput.* 2012. Vol. 41, N 1. P. 219–249.
203. Kartashov A., Frolov A., Goltsev A., Folk R. Quality and efficiency of retrieval for Willshaw-like autoassociative networks: III. Willshaw–Potts model. *Network: Computation in Neural Systems*. 1997. Vol. 8, N 1. P. 71–86.
204. Kasatkina L.M., Kasatkin A.M., Goltsev A.D., Rachkovskij D.A. The implementation of the ideas of academician N.M. Amosov in neural information technologies. *Cybernetics and Computer Engineering*. 2013. Vol. 174. P. 18–29. (in Russian)
205. Kim M., Smaragdis P. Bitwise neural networks. *ICML'15. 2015. Workshop on Resource-Efficient Machine Learning*.
206. Kim S., Choi S. Bilinear random projections for locality-sensitive binary codes. *Proc. CVPR'15*. 2015. P. 1338 – 1346.
207. Kim S., Choi S. Binary embedding with additive homogeneous kernels. *Proc. AAAI'17*. 2017. P. 2094–2100.
208. Kim S., Choi S. Sparse circulant binary embedding: an asymptotic analysis. *IEEE Signal Processing Letters*. 2018. Vol. 25, N 3. P. 432–436.
209. Kim S., Kim J., Choi S. On the optimal bit complexity of circulant binary embedding. *Proc. AAAI'18*. 2018. P. 3423–3430.
210. Klartag B., Mendelson S. Empirical processes and random projections. *Journal of Functional Analysis*. 2005. Vol. 225, N 1. P. 229–245.

211. Kleyko D., Osipov E. On bidirectional transitions between localist and distributed representations: The case of common substrings search using vector symbolic architecture. *Procedia Computer Science*. 2014. Vol. 41. P. 104–113.
212. Kleyko D., Osipov E., Rachkovskij D.A. Modification of holographic graph neuron using sparse distributed representations. *Procedia Computer Science*. 2016. Vol. 88. P. 39–45.
213. Kleyko D., Osipov E., Senior A., Khan A.I., Sekercioglu Y.A. Holographic graph neuron: A bioinspired architecture for pattern processing. *IEEE Trans. Neural Netw. Learn. Syst.* 2017. Vol. 28, N 6. P. 1250–1262.
214. Kleyko D., Rahimi A., Rachkovskij D., Osipov E., Rabaey J. Classification and recall with binary hyperdimensional computing: Tradeoffs in choice of density and mapping characteristics. *IEEE Trans. Neural Netw. Learn. Syst.* 2018. Vol. 29. P. 5880–5898.
215. Knuth D. E. Big omicron and big omega and big theta. *ACM Sigact News*. 1976. Vol. 8, N 2. P. 18–24.
216. Krahermer F., Mendelson S., Rauhut H. Suprema of chaos processes and the Restricted Isometry Property. *Comm. Pure Appl. Math.* 2014. Vol. 67, N 11. P. 1877–1904.
217. Krahermer F., Ward R. New and improved Johnson-Lindenstrauss embeddings via the Restricted Isometry Property. *SIAM J. Math. Anal.* 2011. Vol. 43, N 3. P. 1269–1281.
218. Krahermer F., Ward R. A unified framework for linear dimensionality reduction in L_1 . *Results in Mathematics*. 2015. P.1–23.
219. Kriege N., Neumann M., Kersting K., Mutzel P. Explicit versus implicit graph feature maps: A computational phase transition for walk kernels. *Proc. ICDM'14*. 2014. P. 881–886.
220. Korolev V., Shevtsova I. An improvement of the Berry-Esseen inequality with applications to Poisson and mixed Poisson random sums. *Scandinavian Actuarial Journal*. 2012. Vol. 2012, N 2. P. 81–105.
221. Kulis B. Metric learning: a survey. *Foundations and Trends in Machine Learning*. 2012. Vol. 5, N 4. P. 287–364.
222. Kulis B., Grauman K. Kernelized locality-sensitive hashing. *IEEE Trans. PAMI*. 2012. Vol. 34, N 6. P. 1092–1104.
223. Kumar S., Mohri M., Talwalkar A.. Sampling methods for the nystrom method. *Journal of Machine Learning Research*. 2012. Vol. 13, N 1. P. 981–1006.
224. Kushilevitz E., Nisan N. Communication Complexity. Cambridge: Cambridge University Press, 2006. 208 p.
225. Kushilevitz E., Ostrovsky R., Rabani Y. Efficient search for approximate nearest neighbor in high dimensional spaces. *SIAM Journal on Computing*. 2000. Vol. 30, N 2. P. 457–474.
226. Kussul E., Baidyk T., Kasatkina L. Lukovich V. Rosenblatt perceptrons for handwritten digit recognition. *Proc. IJCNN'01*. Washington, USA. 2001. P. 1516–1521.

227. Kussul E.M., Baidyk T.N., Lukovich V.V., Rachkovskij D.A. Adaptive neural network classifier with multfloat input coding. *Proc. Neuro-Nimes'93*. 1993. P. 209–216.
228. Kussul E.M., Baidyk T.N., Lukovich V.V., Rachkovskij D.A. Adaptive high performance classifier based on random threshold neurons. *Proc. EMCSR'94*. 1994. P. 1687–1694.
229. Kussul E.M., Kasatkina L.M., Rachkovskij D.A., Wunsch D.C. Application of random threshold neural networks for diagnostics of micro machine tool condition. *Proc. IJCNN'98*. 1998. P. 241–244.
230. Kussul E.M., Rachkovskij D.A. Multilevel assembly neural architecture and processing of sequences. In: *Neurocomputers and Attention, II: Connectionism and neurocomputers*. Manchester and New York: Manchester University Press, 1991. P.577–590.
231. Kussul E.M., Rachkovskij D.A., Baidyk T.N. Associative-projective neural networks: architecture, implementation, applications. *Proc. Neuro-Nimes'91*. 1991. P. 463–476.
232. Kussul E.M., Rachkovskij D.A., Baidyk T.N. On image texture recognition by associative-projective neurocomputer. *Proc. ANNIE'91*. 1991. P. 453–458.
233. Kussul E.M., Rachkovskij D.A., Wunsch D.C. The random subspace coarse coding scheme for real-valued vectors. *Proc. IJCNN'99*. 1999. P. 450–455.
234. Kussul N., Lavreniuk M., Shelestov A., Skakun S. Crop inventory at regional scale in Ukraine: developing in season and end of season crop maps with multi-temporal optical and SAR satellite imagery. *European Journal of Remote Sensing*. 2018. Vol. 51, N 1. P. 627–636.
235. Kussul N., Lemoine G., Gallego F. J., Skakun S. V., Lavreniuk M., Shelestov A. Y. Parcel-based crop classification in Ukraine using Landsat-8 data and Sentinel-1A data. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 2016. Vol. 9, N 6. P. 2500–2508.
236. Kussul N.N., Sokolov B.V., Zyelyk Y.I., Zelentsov V.A., Skakun S.V., Shelestov A.Y. Disaster risk assessment based on heterogeneous geospatial information. *J. of Automation and Information Sciences*. 2010. Vol. 42, N 12. P. 32–45.
237. Kyng R. J., Phillips J. M., Venkatasubramanian S. Johnson-Lindenstrauss dimensionality reduction on the simplex. *Proc. FWCG'10*. 2010. P. 1–4.
238. Larsen K.G., Nelson J. The Johnson-Lindenstrauss lemma is optimal for linear dimensionality reduction. *Proc. ICALP'16*. 2016. P. 82:1–82:11.
239. Larsen K.G., Nelson J. Optimality of the Johnson-Lindenstrauss lemma. *Proc. FOCS'17*. 2017. P. 633–638.
240. Le Q., Sarlos T., Smola A.J. Fastfood – Computing hilbert space expansions in loglinear time. *JMLR W&CP*. 2013. Vol. 28, N 3. P. 244–252.
241. Lee J., Mendel M., Naor A. Metric structures in l_1 : Dimension, snowflakes, and average distortion. *European Journal of Combinatorics*. 2005. Vol. 26, N 8. P.1180–1190.

242. Lee J.R., Naor A. Embedding the diamond graph in L_p and dimension reduction in L_1 . *Geometric and Functional Analysis*. 2004. Vol. 14, N 4. P. 745–747.
243. Leutenegger S., Chli M., Siegwart R. Brisk: Binary robust invariant scalable keypoints. *Proc. ICCV'11*. 2011. P. 2548–2555.
244. Li M., Bi W., Kwok J.T., Lu B.-L. Large-scale Nystrom kernel matrix approximation using randomized SVD. *IEEE Transactions on Neural Networks and Learning Systems*. 2015. Vol. 26, N 1. P. 152–164.
245. Li P. Very sparse stable random projections, estimators and tail bounds for stable random projections. *Journal of Machine Learning Research*. 2007. Vol. 8. P. 2497–2532.
246. Li P. Very sparse stable random projections for dimension reduction in l_α ($0 < \alpha \leq 2$) norm. *Proc. KDD'07*. 2007. P.440–449.
247. Li P. Estimators and tail bounds for dimension reduction in l_α ($0 < \alpha \leq 2$) using stable random projections. *Proc. SODA'08*. 2008. P. 10–19.
248. Li P. Computationally efficient estimators for dimension reductions using stable random projections. *Proc. ICDM'08*. 2008. P. 403–412.
249. Li P. Sign stable random projections for large-scale learning. *arXiv:1504.07235*. 27 Apr 2015.
250. Li P. Linearized GMM kernels and normalized Random Fourier Features. *Proc. KDD '17*. 2017. P. 315–324
251. Li P. Sign-full random projections. *arXiv.org:1805.00533*. 26 Apr 2018.
252. Li P., Church K.W. A sketch algorithm for estimating two-way and multi-way associations. *Computational Linguistics*. 2007. Vol. 33, N 3. P. 305–354.
253. Li P., Church K.W., Hastie T.J. Conditional random sampling: A sketch-based sampling technique for sparse data. *Proc. NIPS'06*. 2006. P. 873–880.
254. Li P., Church K.W., Hastie T.J. One sketch for all: Theory and applications of conditional random sampling. *Proc. NIPS'08*. 2008. P. 953–960.
255. Li P., Hastie T.J. A unified near-optimal estimator for dimension reduction in l_α ($0 < \alpha \leq 2$) using stable random projections. *Proc. NIPS'07*. 2007. P. 905–912.
256. Li P., Hastie T.J., Church K.W. Very sparse random projections. *Proc. KDD'06*. 2006. P. 287–296.
257. Li P., Hastie T.J., Church K.W. Improving random projections using marginal information. *Proc. COLT'06*. 2006. P. 635–649.
258. Li P., Hastie T.J., Church K.W. Nonlinear estimators and tail bounds for dimensional reduction in l_1 using Cauchy random projections. *Journal of Machine Learning Research*. 2007. Vol. 8. P. 2497–2532.
259. Li P., Konig A.C. Theory and applications of b-bit minwise hashing. *Communications of the ACM*. 2011. Vol. 54, N 8. P. 101–109.
260. Li P., Mahoney M.W., She Y. Approximating higher-order distances using random projections. *Proc. UAI'10*. 2010. P. 312–321.
261. Li P., Mitzenmacher M., Shrivastava A. Coding for random projections. *Proc. ICML'14*. 2014. P. 676–684.

262. Li P., Mitzenmacher M., Shrivastava A. 2-Bit random projections, nonlinear estimators, and approximate near neighbor search. *arXiv:1602.06577*. 21 Feb 2016.
263. Li P., Mitzenmacher M., Slawski M. Quantized random projections and non-linear estimation of cosine similarity. *Proc. NIPS'16*. 2016. P. 2748–2756.
264. Li P., Samorodnitsky G., Hopcroft J. Sign cauchy projections and chi-square kernel. *Proc. NIPS'13*. 2013. P. 2571–2579.
265. Li P., Shrivastava A., Moore J.L., Konig A.C. Hashing algorithms for large-scale learning. *Proc. NIPS'11*. 2011. P. 2672–2680.
266. Li P., Slawski M. Simple strategies for recovering inner products from coarsely quantized random projections. *Proc. NIPS'17*. 2017. P. 4570–4579.
267. Li P., Zhang C.-H. Theory of the GMM kernel. *Proc. WWW'17*. 2017. P. 1053–1062.
268. Liberty E., Ailon N., Singer A. Dense fast random projections and lean walsh transforms. *Discrete and Computational Geometry*. 2011. Vol. 45, N 1. P.34–44.
269. Liberty E., Zucker S.W. The mailman algorithm: A note on matrix–vector multiplication. *Inf. Process. Lett.* 2009. Vol. 109, N 3. P. 179–182.
270. Linial N. Finite metric spaces — combinatorics, geometry and algorithms. *Proc. of the International Congress of Mathematicians*. 2002. Vol. III. P. 573–586.
271. Linial N., London E., Rabinovich Y. The geometry of graphs and some of its algorithmic implications. *Combinatorica*. 1995. Vol. 15, N 2. P. 215–245.
272. Lv Q., Charikar M., Li K. Image similarity search with compact data structures. *Proc. CIKM'04*. 2004. P. 208–217.
273. Lowe D.G. Object recognition from local scale-invariant features. *Proc. ICCV'99*. 1999. P. 1150–1157.
274. Lowe D. Towards a computational model for object recognition in IT cortex. *Proc. BMCV'00*. 2000. P. 20–31.
275. Lowe D. Distinctive image features from scale-invariant keypoints. *IJCV*. 2004. Vol. 60, N 2. P. 91–110.
276. Lukovich V.V., Goltsev A.D., Rachkovskij D.A. Neural network classifiers for micromechanical equipment diagnostics and micromechanical product quality inspection. *Proc. EUFIT'97*. 1997. P. 534–536.
277. Magen A. Dimensionality reductions in ℓ_2 that preserve volumes and distance to affine spaces. *Discrete Comput. Geom.* 2007. Vol. 38, N 1. P.139–153.
278. Mahoney M.W., Drineas P. CUR matrix decompositions for improved data analysis. *Proc. Natl. Acad. Sci. USA*. 2009. Vol. 106. P. 697–702.
279. Mahoney M.W. Randomized algorithms for matrices and data. *Foundations and Trends in Machine Learning*. 2011. Vol. 3, N 2. P. 123–224.
280. Manku G.S., Jain A., Sarma A.D. Detecting near-duplicates for web crawling. *Proc. WWW'07*. 2007. P. 141–150.
281. Martinetz T., Labusch K., Schneegass D. SoftDoubleMaxMinOver: Perceptron-like training of Support Vector Machines. *IEEE Transactions on Neural Networks*. 2009. Vol. 20, N 7. P. 1061–1072.

282. Matousek J. On the distortion required for embedding finite metric spaces into normed spaces. *Israel J. Math.* 1996. Vol. 93. P. 333–344.
283. Matousek J. On variants of the Johnson Lindenstrauss lemma. *Random Structures and Algorithms.* 2008. Vol. 33, N 2. P. 142–156.
284. Matousek J. Lecture notes on metric embeddings. 2013. 126 p.
285. Matousek J., Naor A. 2011. Open problems on embeddings of finite metric spaces. 2011. <http://students.mimuw.edu.pl/~mk249015/www10/metrop.pdf>
286. McMahan H. B., G. Holt, D. Sculley, M. Young, D. Ebner, J. Grady, L. Nie, T. Phillips, E. Davydov, D. Golovin, S. Chikkerur, D. Liu, M. Wattenberg, A.M. Hrafnkelsson, T. Boulos, J. Kubica. Ad click prediction: a view from the trenches. *Proc. KDD'13.* 2013. P. 1222–1230.
287. Min K., Yang L., Wright J., Wu L., Hua X.-S., Ma Y. Compact projection: Simple and efficient near neighbor search with practical memory requirements. *Proc. CVPR'10.* 2010. P. 3477–3484.
288. Misuno I.S., Rachkovskij D.A., Slipchenko S.V. Vector and distributed representations reflecting semantic relatedness of words. *Mathematical machines and systems.* 2005. N 3. P. 50–67. (in Russian)
289. Misuno I.S., Rachkovskij D.A., Slipchenko S.V., Sokolov A.M. Searching for text information with the help of vector representations. *Problems of Programming.* 2005. N 4. P. 50–59. (in Russian)
290. Mitzenmacher M., Pagh R., Pham N. Efficient estimation for high similarities using odd sketches. *Proc. WWW'14.* 2014. P. 109–118.
291. Moczulski M., Denil M., Appleyard J., de Freitas N. Acdc: A structured efficient linear layer. *ICLR'16.* 2016.
292. Motwani R., Raghavan P. Randomized Algorithms. New York: Cambridge University Press, 1995. 496 p.
293. Mu Y., Hua G., Fan W., Chang S.-F. Hash-SVM: scalable kernel machines for large-scale visual classification. *Proc. CVPR'14.* 2014. P. 446–451.
294. Mu Y., Yan S. Non-Metric Locality Sensitive Hashing. *Proc. AAAI'10.* 2010. P. 539–544.
295. Musco C., Musco C. Recursive sampling for the Nyström method. *Proc. NIPS'17.* 2017. P. 3833–3845.
296. Myleiko G.L., Pereverzyev S. Jr, Solodky S.G. Regularized Nystrom subsampling in regression and ranking problems under general smoothness assumptions. *Analysis and Applications.*
297. Nagaev S.V., Chebotarev V.I. On the bound of proximity of the binomial distribution to the normal one. *Doklady Mathematics.* 2011. Vol. 83, N 1. P. 19–21.
298. Nakarmi U., Rahnavard N., BCS: compressive sensing for binary sparse signals. *Proc. MILCOM'12.* 2012. P. 1–5.
299. Narayanan S., Nelson J. Optimal terminal dimensionality reduction in Euclidean space. *arXiv:1810.09250.* 22 Oct 2018.
300. Nelson J., Nguyen H.L. Sparsity lower bounds for dimensionality reducing maps. *Proc. STOC'13.* 2013. P. 101–110.

301. Nelson J., Price E., Wootters M. New constructions of RIP matrices with fast multiplication and fewer rows. *Proc. SODA'14*. 2014. P. 1515–1528.
302. Nelson J., Woodruff D. P. Fast Manhattan sketches in data streams. *Proc. PODS'10*. 2010. P. 99–110.
303. Newman I., Rabinovich Y. Finite volume spaces and sparsification. *arXiv:1002.3541*. 2 Aug 2010.
304. Ng J. Y.-H., Yang F., Davis L. S. Exploiting local features from deep networks for image retrieval. *Proc. CVPRW'15*. 2015. P. 53–61.
305. Nisan N. $RL \subseteq SC$. *Proc. STOC'92*. 1992. P. 619–623.
306. Nolan J.P. Stable distributions — models for heavy tailed data. Boston: Birkhauser, 2015. Chapter 1. P. 3–22.
307. Nowicki D., Siegelmann H. Flexible kernel memory. *PLoS ONE*. 2010. 5(6): e10955.
308. Nowicki D., Verga P., Siegelmann H. Modeling Reconsolidation in kernel associative memory. *PLoS ONE*. 2013. 8(8): e68189.
309. Oehlert G.W. A Note on the delta method. *The American Statistician*. 1992. Vol. 46, N 1. P. 27–29.
310. Oliva A., Torralba A. Modeling the shape of the scene: a holistic representation of the spatial envelope. *International Journal of Computer Vision*. 2001. Vol. 42, N 3. P. 145–175.
311. Olshausen B.A., Field D.J. Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.* 2004. Vol. 14. P. 481–487.
312. Oymak S., Thrampoulidis C., Hassibi B. Near-optimal sample complexity bounds for circulant binary embedding. *Proc. ICASSP'17*. 2017. P. 6359–6363.
313. Oymak S., Recht B. Near optimal bounds for binary embeddings of arbitrary sets. *arXiv:1512.04433*. 14 Dec 2015.
314. Oymak S., Recht B., Soltanolkotabi M. Isometric sketching of any set via the Restricted Isometry Property. *Information & Inference*.
315. Page M. Connectionist modelling in psychology: A localist manifesto. *Behavioral and Brain Sciences*. 2000. Vol. 23. P. 443–512.
316. Pan V.Y. Fast feasible and unfeasible matrix multiplication. *arXiv:1804.04102*. 11 Apr 2018.
317. Paschou P., Ziv E., Burchard E.G., Choudhry S., Rodriguez-Cintron W., Mahoney M.W., Drineas P. PCA-correlated SNPs for structure identification in worldwide human populations. *PLoS Genetics*. 2007. Vol. 3. P. 1672–1686.
318. Pasunuri R., China Venkaiah V., Dhariyal B. Ascending and descending order of random projections: comparative analysis of high-dimensional data clustering. *Advances in Intelligent Systems and Computing*. 2019. Vol. 741. P. 133–142.
319. Pauleve L., Jegou H., Amsaleg L. Locality sensitive hashing: A comparison of hash function types and querying mechanisms. *Pattern Recognit. Lett.* 2010. Vol. 31, N 11. P. 1348–1358.

320. Pavlov D.N., Mannila H., Smyth P. Beyond independence: probabilistic models for query approximation on binary transaction data. *IEEE TKDE*. 2003. Vol. 15, N. 6. P. 1409–1421.
321. Pennington J., Yu F.X., Kumar S. Spherical random features for polynomial kernels. *Proc. NIPS'15*. 2015. P. 1846–1854.
322. Perronnin F., Liu Y., Sanchez J., Poirier H. Large-scale image retrieval with compressed fisher vectors. *Proc. CVPR'10*. 2010. P. 3384–3391.
323. Perronnin F., Sanchez J., Mensink T. Improving the fisher kernel for large-scale image classification. *Proc. ECCV'10*. 2010. P. 143–156.
324. Pham N., Pagh R. Fast and scalable polynomial kernels via explicit feature maps. *Proc. KDD'13*. 2013. P. 239–247.
325. Plan Y., Vershynin R. One-bit compressed sensing by linear programming. *Communications on Pure and Applied Mathematics*. 2013. Vol. 66, N. 8. P. 1275–1297.
326. Plan Y., Vershynin R. Dimension reduction by random hyperplane tessellations. *Discrete and Computational Geometry*. 2014. Vol. 51, N 2. P. 438–461.
327. Plate T. Holographic Reduced Representation: Distributed Representation for Cognitive Structures. Stanford: CSLI Publications, 2003. 300 p.
328. Platt J.C. FastMap, MetricMap, and Landmark MDS are all Nystrom algorithms. *Proc. AISTATS'05*. 2005. P. 261–268.
329. Pylieva H., Chernodub A., Grabar N., Hamon T. Improving automatic categorization of technical vs. laymen medical words using fasttext word embeddings. *Proc. IDDM'18*. 2018. P. 93–102.
330. Rachkovskij D.A. Representation and processing of structures with binary sparse distributed codes. *IEEE Transactions on Knowledge and Data Engineering*. 2001. Vol. 13, N 2. P. 261–276.
331. Rachkovskij D.A. Some approaches to analogical mapping with structure sensitive distributed representations. *J. Experimental and Theoretical Artificial Intelligence*. 2004. Vol. 16, N 3. P. 125–145.
332. Rachkovskij D.A. Vector data transformation using random binary matrices. *Cybernetics and Systems Analysis*. 2014. Vol. 50, N 6. P. 960–968.
333. Rachkovskij D.A. Formation of similarity-reflecting binary vectors with random binary projections. *Cybernetics and Systems Analysis*. 2015. Vol. 51, N 2. P. 313–323.
334. Rachkovskij D.A. Estimation of vectors similarity by their randomized binary projections. *Cybernetics and Systems Analysis*. 2015. Vol. 51, N 5. P. 808–818.
335. Rachkovskij D.A. Real-valued vectors for fast distance and similarity estimation. *Cybernetics and Systems Analysis*. 2016. Vol. 52, N 6. P. 967–988.
336. Rachkovskij D.A. Binary vectors for fast distance and similarity estimation. *Cybernetics and Systems Analysis*. 2017. Vol. 53, N 1. P. 138–156.
337. Rachkovskij D.A. Distance-based index structures for fast similarity search. *Cybernetics and Systems Analysis*. 2017. Vol. 53, N 4. P. 636–658.

338. Rachkovskij D.A. Index structures for fast similarity search for binary vectors. *Cybernetics and Systems Analysis*. 2017. Vol. 53, N 5. P. 799–820.
339. Rachkovskij D.A. Index structures for fast similarity search for real-valued vectors. I. *Cybernetics and Systems Analysis*. 2018. Vol. 54, N 1. P. 152–164.
340. Rachkovskij D.A. Index structures for fast similarity search for real-valued vectors. II. *Cybernetics and Systems Analysis*. 2018. Vol. 54, N 2. P. 320–335.
341. Rachkovskij D.A., Kussul E.M. Binding and normalization of binary sparse distributed representations by context-dependent thinning. *Neural Computation*. 2001. Vol. 13, N 2. P. 411–452.
342. Rachkovskij D.A., Kussul E.M., Baidyk T.N. Building a world model with structure-sensitive sparse binary distributed representations. *BICA*. 2013. Vol. 3. P. 64–86.
343. Rachkovskij D.A., Misuno I.S., Slipchenko S.V. Randomized projective methods for construction of binary sparse vector representations. *Cybernetics and Systems Analysis*. 2012. Vol. 48, N 1. P. 146–156.
344. Rachkovskij D.A., Revunova E.G. Intelligent gamma-ray data processing for environmental monitoring. In: *Intelligent Data Processing in Global Monitoring for Environment and Security*. Kiev-Sofia: ITHEA, 2011. P. 136–157.
345. Rachkovskij D.A., Revunova E.G. Randomized method for solving discrete ill-posed problems. *Cybernetics and Systems Analysis*. 2012. Vol. 48, N. 4. P. 621–635.
346. Rachkovskij D., Slipchenko S. Similarity-based retrieval with structure-sensitive sparse binary distributed representations. *Computational Intelligence*. 2012. Vol. 28, N 1. P. 106–129.
347. Rachkovskij D.A., Slipchenko S.V., Frolov A. A., Gusek D. Resolution of binary coding of real-valued vectors by hyperrectangular receptive fields. *Cybernetics and Systems Analysis*. 2005. Vol. 41, N 5. P. 635–646.
348. Rachkovskij D.A., Slipchenko S.V., Kussul E.M., Baidyk T.N. A binding procedure for distributed binary data representations. *Cybernetics and Systems Analysis*. 2005. Vol. 41, N 3. P. 319–331.
349. Rachkovskij D.A., Slipchenko S.V., Kussul E.M., Baidyk T.N. Properties of numeric codes for the scheme of random subspaces RSC. *Cybernetics and Systems Analysis*. 2005. Vol. 41, N 4. P. 509–520.
350. Rachkovskij D.A., Slipchenko S.V., Kussul E.M., Baidyk T.N. Sparse binary distributed encoding of scalars. *Journal of Automation and Information Sciences*. 2005. Vol. 37, N 6. P. 12–23.
351. Rachkovskij D.A., Slipchenko S.V., Misuno I.S., Kussul E.M., Baidyk T.N. Sparse binary distributed encoding of numeric vectors. *J. of Automation and Inf. Sci.* 2005. Vol. 37, N 11. P. 47–61.
352. Rahimi A., Recht B. Random features for large-scale kernel machines. *Proc. NIPS'07*. 2007. P.1177–1184.
353. Raginsky M., Lazebnik S. Locality-sensitive binary codes from shift invariant kernels. *Proc. NIPS'09*. 2009. P. 1509–1517.

354. Rajan K., Abbot L.F. Eigenvalue spectra of random matrices for neural networks. *Physical Review Letters*. 2006. Vol. 97:188104.
355. Rasanen O.J., Saarinen J.P. Sequence prediction with sparse distributed hyperdimensional coding applied to the analysis of mobile phone use patterns. *IEEE Trans. Neural Netw. Learn. Syst.* 2016. Vol. 27, N 9. P. 1878–1889.
356. Rastegari M., Ordonez V., Redmon J., Farhadi A. Xnor-net: Imagenet classification using binary convolutional neural networks. *Proc. ECCV'16*. 2016. P. 525–542.
357. Rauhut H., Romberg J., Tropp J. Restricted isometries for partial random circulant matrices. *Applied and Computational Harmonic Analysis*. 2012. Vol. 32, N 2. P. 242–254.
358. Recchia G., Sahlgren M., Kanerva P., Jones M. Encoding sequential information in semantic space models: Comparing holographic reduced representation and random permutation. *Comput. Intell. Neurosci.* 2015. Vol. 2015:986574. 18 p.
359. Ren S., Cao X., Wei Y., Sun J. Face alignment at 3000 fps via regressing local binary features. *Proc. CVPR'14*. 2014. P. 1685–1692.
360. Revunova E.G. Randomization approach to the reconstruction of signals resulted from indirect measurements. *Proc. ICIM'13*. 2013. P. 203–208.
361. Revunova E.G. Analytical study of the error components for the solution of discrete ill-posed problems using random projections. *Cybernetics and Systems Analysis*. 2015. Vol. 51, N. 6. P. 978–991.
362. Revunova E.G. Model selection criteria for a linear model to solve discrete ill-posed problems on the basis of singular decomposition and random projection. *Cybernetics and Systems Analysis*. 2016. Vol. 52, N.4. P. 647–664.
363. Revunova E.G. Averaging over matrices in solving discrete ill-posed problems on the basis of random projection. *Proc. CSIT'17*. 2017. Vol. 1. P. 473 – 478.
364. Revunova E.G. Solution of the discrete ill-posed problem on the basis of singular value decomposition and random projection. *Advances in Intelligent Systems and Computing II*. Cham: Springer, 2018. P. 434–449.
365. Revunova E.G. Increasing the accuracy of the solution of discrete ill-posed problems by the method of random projections. *Control systems and machines*. 2018. N 1. P. 16–27. (in Ukrainian)
366. Revunova E.G. Improving the accuracy of the solution of discrete ill-posed problem by random projection. *Cybernetics and Systems Analysis*. 2018. Vol. 54, N 5. P. 842–852.
367. Revunova E.G., Rachkovskij D.A. Using randomized algorithms for solving discrete ill-posed problems. *Intern. Journal Information Theories and Applications*. 2009. Vol. 16, N 2. P. 176–192.
368. Revunova E.G., Rachkovskij D.A. Random projection and truncated SVD for estimating direction of arrival in antenna array. *Cybernetics and Computer Engineering*. 2018. N 3(193). P. 5–26.

369. Revunova E.G., Tyshchuk A.V. A model selection criterion for solution of discrete ill-posed problems based on the singular value decomposition. *Proc. IWIM'2015*. 2015. P.43–47.
370. Reznik A.M., Sitchov A.S., Dekhtyarenko O.K., Nowicki D.W. Associative memories with killed neurons: the methods of recovery. *Proc. IJCNN'03*. 2003. P. 2579–2582.
371. Reznik A.M., Shirshov Yu. M., Snopok B.A., Nowicki D.W., Dekhtyarenko A.K., Kruglenko I.V. Associative memories for chemical sensing. *Proc. ICONIP'02*. 2002. P. 2630–2634.
372. Riesen K., Neuhaus M., Bunke H. Graph embedding in vector spaces by means of prototype selection. *Proc. GbRPR'07*. 2007. P. 383–393.
373. Rublee E., Rabaud V., Konolige K., Bradski G. Orb: An efficient alternative to SIFT or SURF. *Proc. ICCV'11*. 2011. P. 2564–2571.
374. Rudi A., Camoriano R., Rosasco L. Less is more: Nystrom computational regularization. *Proc. NIPS'15*. 2015. P. 1657–1665.
375. Samet H. Foundations of Multidimensional and Metric Data Structures. San Francisco: Morgan Kaufmann, 2006. 1024 p.
376. Sankowski P., Wygocki P. Approximate nearest neighbors search without false negatives for l_2 for $c > \sqrt{\log \log \{n\}}$. *Proc. ISAAC'17*. 2017. P. 63:1–63:12.
377. Schechtman G. Random embeddings of Euclidean spaces in sequence spaces. *Israel Journal of Mathematics*. 1981. Vol. 40, N 2. P. 187–192.
378. Schechtman G., More on embedding subspaces of L_p in ℓ_r^n . *Compositio Math*. 1987. Vol. 61, N 2. P. 159–169.
379. Schechtman G. Two observations regarding embedding subsets of Euclidean spaces in normed spaces. *Adv. Math*. 2006. Vol. 200, N 1. P. 125–135.
380. Schechtman G. Dimension reduction in l_p , $0 < p < 2$. *arXiv:1110.2148*. 10 Oct 2011.
381. Schlesinger M.I., Vodolazskii Y.V. and Yakovenko V.M. Recognizing the similarity of polygons in a strengthened Hausdorff metric. *Cybern. Syst. Anal*. 2014. Vol. 50, N 3. P. 476–486.
382. Schlesinger M.I., Vodolazskiy E.V., Yakovenko V.M. Frechet similarity of closed polygonal curves. *Intern. J. of Computational Geometry & Applications*. 2016. Vol. 26, N 1. P. 53–66.
383. Scholkopf B., Smola A.J. Learning With Kernels, Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, 2001. 648 p.
384. Scholkopf B., Smola A., Muller K.-R. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*. 1998. Vol. 10, N 5. P. 1299–1319.
385. Shalev-Shwartz S., Singer Y., Srebro N. Pegasos: primal estimated sub-gradient solver for SVM. *Proc. ICML'07*. 2007. P. 807–814.
386. Shen F., Mu Y., Yang Y., Liu W., Liu L., Song J., Shen H.T. Classification by retrieval: Binarizing data and classifiers. *Proc. SIGIR'17*. 2017. P. 595–604.
387. Shervashidze N., Vishwanathan S. V. N., Petri T., Mehlhorn K., Borgwardt K. Efficient graphlet kernels for large graph comparison. *JMLR: W&CP*. 2009. Vol. 5. P. 488–495.

388. Shevtsova I.G. On the absolute constants in the Berry–Esseen inequality and its structural and nonuniform improvements. *Inform. Primen.* 2013. Vol. 7, N 1. P. 124–125.
389. Shevtsova I.G. On the absolute constants in the Berry–Esseen-type inequalities. *Doklady Mathematics.* 2014. Vol. 89, N 3. P. 378–381.
390. Shrivastava A., Li P. Fast near neighbor search in high-dimensional binary data. *Proc. ECML PKDD'12.* 2012. P. 474–489.
391. Shrivastava A., Li P. In defense of MinHash over SimHash. *Proc. AISTATS'14.* 2014. P. 886–894.
392. Shrivastava A., Li P. Asymmetric minwise hashing for indexing binary inner products and set containment. *Proc. WWW'15.* 2015. P. 981–991.
393. Si S., Hsieh C.-J., Dhillon I. Memory efficient kernel approximation. *Proc. ICML'14.* 2014. P. 701–709.
394. Sivertsen J. Fast nearest neighbor preserving embeddings. *arXiv:1707.06867.* 21 Jul 2017.
395. Sindhwani V., Sainath T., Kumar S. Structured transforms for smallfootprint deep learning. *Proc. NIPS'15.* 2015. P. 3070–3078.
396. Sivakumar D. Algorithmic derandomization via complexity theory. *Proc. STOC'02.* 2002. P. 619–626.
397. Sivic J., Zisserman A. Video Google: A text retrieval approach to object matching in videos. *Proc. ICCV'03.* 2003. P. 1470–1477.
398. Slipchenko S.V., Misuno I.S., Rachkovskij D.A. Properties of coarse coding with random hyperrectangle receptive fields. *Mathematical machines and systems.* 2005. N 4. P. 15–29. (in Russian)
399. Slipchenko S.V., Rachkovskij D.A. Analogical mapping using similarity of binary distributed representations. *Int. J. Information Theories and Applications.* 2009. Vol. 16, N 3. P. 269–290.
400. Slipchenko S. V., Rachkovskij D. A., Misuno I. S. Decoding binary distributed representations of numerical vectors. *Computer Mathematics.* 2005. N 3. P. 108–120. (in Russian)
401. Sokolov A. Vector representations for efficient comparison and search for similar strings. *Cybernetics and System Analysis.* 2007. Vol. 43, N 4. P.484–498.
402. Sokolov A. and Rachkovskij D. Approaches to sequence similarity representation. *J. Information Theories and Applications.* 2005. Vol.13, N 3. P. 272–278.
403. Sriperumbudur B. K., Szabo Z. Optimal rates for random Fourier features. *Proc. NIPS'15.* 2015. P. 1144–1152.
404. Steinwart I., Christmann A. Support Vector Machines. New York: Springer, 2008. 601 p.
405. Sun F., Guo J., Lan Y., Xu J., Cheng X. Sparse word embeddings using l1 regularized online learning. *Proc. IJCAI'16.* 2016. P. 2915–2921.
406. Sun S., Zhao J., Zhu J. A review of Nystrom methods for large-scale machine learning. *Information Fusion.* 2015. Vol. 26(C). P. 36–48.
407. Sutherland D. J., Schneider J. On the error of random Fourier features. *Proc. UAI'15.* 2015. P. 862–871.

408. Tang W., Hua G., Wang L. How to train a compact binary neural network with high accuracy? *Proc. AAAI'17*. 2017. P. 2625–2631.
409. Thorpe S. Localized versus distributed representations. In: *The Handbook of Brain Theory and Neural Networks*. Edited by M. A. Arbib. Cambridge, MA: The MIT Press, 2003. P. 643–646.
410. Thorup M. Fast and powerful hashing using tabulation. *Communications of the ACM*. 2017. Vol. 60, N 7. P. 94–101.
411. Tropp J.A. An introduction to matrix concentration inequalities. *Foundations and Trends in Machine Learning*. 2015. Vol. 8, N 1-2. P. 1–230.
412. Tyurin I. S. A refinement of the remainder in the Lyapunov theorem. *Theory of Probab. and its Applications*. 2012. Vol. 56, N 4. P. 693–696.
413. Uchaikin V.V., Zolotarev V.M. Chance and Stability. *Stable Distributions and their Applications*. VSP, 1999. 596 p.
414. Van der Maaten L. J. P., Postma E. O., Van den Herik H. J. Dimensionality reduction: A comparative review. TiCC-TR 2009-005. Tilburg University Technical Report, 2009.
415. Vedaldi A., Zisserman A.. Efficient additive kernels via explicit feature maps. *IEEE trans. PAMI*. 2012. Vol. 34, N 3. P. 480–492.
416. Vempala S.S. *The Random Projection Method*. American Math. Soc., 2004. 105 p.
417. Ventkatasubramanian S., Wang Q. The Johnson-Lindenstaus transform: An empirical study. *Proc. ALLENEX'11*. 2011. P. 164–173.
418. Vershynin R. Introduction to the non-asymptotic analysis of random matrices. In: *Compressed Sensing, Theory and Applications*. Cambridge, Cambridge University Press, 2012. P. 210–268.
419. Vybiral J. A variant of the Johnson–Lindenstrauss lemma for circulant matrices. *Journal of Functional Analysis*. 2011. Vol. 260, N 4. P. 1096–1105.
420. Uchida Y. Local feature detectors, descriptors, and image representations: a survey. *arXiv:1607.08368*. 28 Jul 2016.
421. Walck C. Hand-book on statistical distributions for experimentalists. Internal Report SUF-PFY/96-01 (last modification 10 Sept. 2007). Fysikum University of Stockholm, Particle Physics Group, 2007. 190 p.
422. Wang D., Zhou Q., Hussain A. Deep and sparse learning in speech and language processing: an overview. *Proc. BICS'16*. 2016: P. 171–183.
423. Wang J., Liu W., Kumar S., Chang S.-F. Learning to hash for indexing big data: A survey. *Proceedings of the IEEE*. 2016. Vol. 104, N 1. P. 34–57.
424. Wang J., Shen H.T., Song J., Ji J. Hashing for similarity search: A survey. *arXiv:1408.2927*. 13 Aug 2014.
425. Wang J., Zhang T., Song J., Sebe N., Shen H.T. A survey on learning to hash. *IEEE trans. PAMI*. 2018. Vol. 40, N 4. P. 769–790.
426. Wang S., Luo L., Zhang Z. SPSD matrix approximation via column selection: theories, algorithms, and extensions. *Journal of Machine Learning Research*. 2016. Vol. 17, N 1. P.1697–1745.

427. Wang Z., Dong W., Josephson W., Lv Q., Charikar M., Li K. Sizing sketches: a rank-based analysis for similarity search. *Proc. SIGMETRICS'07*. 2007. P. 157–168.
428. Wei Y., Xie P., Zhang L. Tikhonov regularization and randomized GSVD. *SIAM J. Matrix Anal. Appl.* 2016. Vol. 37, N 2. P. 649–675.
429. Weinberger K., Dasgupta A., Langford J., Smola A., Attenberg J. Feature hashing for large scale multitask learning. *Proc. ICML'09*. 2009. P. 1113–1120.
430. Williams C. K. I., Seeger M. Using the Nystrom method to speed up kernel machines. *Proc. NIPS'00*. 2000. P. 682–688.
431. Woodruff D.P. Sketching as a tool for numerical linear algebra. *Foundations and Trends in Theoretical Computer Science*. 2014. Vol. 10, N 1–2. P. 1–157.
432. Xia H., Wu P., Hoi S. C., Jin R. Boosting multi-kernel localitysensitive hashing for scalable image retrieval. *Proc. SIGIR'12*. 2012. P. 55–64.
433. Xia Y., He K., Kohli P., Sun J. Sparse projections for high-dimensional binary codes. *Proc. CVPR'15*. 2015. P. 3332–3339.
434. Xiang H., Zou J. Regularization with randomized SVD for large-scale discrete inverse problems. *Inverse Problems*. 2013. Vol. 29, N 8: 085008. P. 1–21.
435. Xiang H., Zou J. Randomized algorithms for large-scale inverse problems with general Tikhonov regularizations. *Inverse Problems*. 2015. Vol. 31, N 8: 085008. P. 1–24.
436. Yang T., Li Y.-F., Mahdavi M., Jin R., Zhou Z.-H. Nystrom method vs random Fourier features: A theoretical and empirical comparison. *Proc. NIPS'12*. 2012. P. 485–493.
437. Yang Z., Moczulski M., Denil M., de Freitas N., Smola A., Song L., Wang Z. Deep fried convnets. *Proc. ICCV'15*. 2015. P. 1476–1483.
438. Yang Y., Pilanci M., Wainwright M.J. Randomized sketches for kernels: Fast and optimal non-parametric regression. *Annals of Statistics*. 2017. Vol. 45, N 3. P. 991–1023.
439. Yang J., Sindhvani V., Avron H., Mahoney M.W. Quasi-Monte Carlo feature maps for shift-invariant kernels. *Proc. ICML'14*. 2014. P. 485–493.
440. Yang J., Sindhvani V., Fan Q., Avron H., Mahoney, M.W. Random laplace feature maps for semigroup kernels on histograms. *Proc. CVPR'14*. 2014. P. 971–978.
441. Yi X., Caramanis C., Price E. Binary embedding: Fundamental limits and fast algorithm. *JMLR: W&CP*. 2015. Vol. 37. P. 2162–2170.
442. Yip C.-W., Mahoney M.W., Szalay A.S., Csabai I., Budavari T., Wyse R.F.G., Dobos L. Objective identification of informative wavelength regions in galaxy spectra. *The Astronomical Journal*. 2014. Vol. 147, N 5:110.
443. Yu F.X., Bhaskara A., Kumar S., Gong Y., Chang S.-F. On binary embedding using circulant matrices. *Journal of Machine Learning Research*. 2018. Vol. 18(150). P. 1–30.
444. Yu F.X., Kumar S., Gong Y., Chang S.-F. Circulant binary embedding. *Proc. ICML'14*. 2014. P. 946–954.

445. Yu F.X., Kumar S., Rowley H., Chang S.-F. Compact nonlinear maps and circulant extensions. *arXiv:1503.03893*. 12 Mar 2015.
446. Yu F., Suresh A., Choromanski K., Holtmann-Rice D., Kumar S. Orthogonal random features. *Proc. NIPS'16*. 2016. P. 1083–1991.
447. Zaychenko Y., Bodyanskiy Y., Tyshchenko O., Boiko O., Hamidov G. Hybrid GMDH-neuro-fuzzy system and its training scheme. *International Journal of Information Theories and Applications*. 2018. Vol. 25, N 1. P. 18–33.
448. Zezula P., Amato G., Dohnal V., Batko M. Similarity Search: The Metric Space Approach. New York: Springer, 2006. 220 p.
449. Zhang H., Cheng L. New bounds for circulant Johnson-Lindenstrauss embeddings. *Communications in Mathematical Sciences*. 2014. Vol. 12, N 4. P. 695–705.
450. Zhang H., Zhang Q. Embedjoin: Efficient edit similarity joins via embeddings. *Proc. KDD'17*. 2017. P. 585–594.
451. Zhang L., Wei Y. Randomized core reduction for discrete ill-posed problem. *arXiv:1808.02654*. 8 Aug 2018.
452. Zhang X., Yu F.X., Guo R., Kumar S., Wang S., Chang S.-F. Fast orthogonal projection based on Kronecker product. *Proc. ICCV'15*. 2015. P. 2929–2937.
453. Zhao D., Lin Z., Tang X. Laplacian PCA and its applications. *Proc. ICCV'07*. 2007. P. 1–8.
454. Zolotarev V.M. One-dimensional Stable Distributions. Translations of Mathematical Monographs. Vol. 65. American Mathematical Society, 1986. 284 p.
455. Zolotukhin A., Nagaev S., Chebotarev V. On a bound of the absolute constant in the Berry–Esseen inequality for i.i.d. Bernoulli random variables. *Modern Stochastics: Theory and Applications*. 2018. Vol. 5, N 3. P. 385–410.
456. Амосов Н.М., Байдык Т.Н., Гольцев А.Д., Касаткин А.М., Касаткина Л.М., Кусскуль Э.М., Рачковский Д.А. Нейрокомпьютеры и интеллектуальные роботы. Киев: Наукова думка, 1991. 272 с.
457. Амосов Н.М., Касаткин А.М., Касаткина Л.М., Талаев С.А. Автоматы и разумное поведение. Киев: Наукова думка, 1973. 374 с.
458. Антомонов М.Ю. Математическая обработка и анализ медико-биологических данных. 2-е изд. Киев: МИЦ «Мединформ», 2018. 558 с.
459. Байдык Т.Н. Нейронные сети и задачи искусственного интеллекта. Киев: Наукова думка, 2001. 263 с.
460. Вентцель Е.С. Теория вероятностей. М.: Наука, 1969. 576 с.
461. Гольцев А.Д., Гриценко В.И. Нейросетевые технологии в задаче распознавания рукописных символов. *УСiМ*. 2018. N 4. С. 3–20.
462. Гриценко В.І., Рачковський Д.А., Ревунова О.Г. Спосіб визначення фізичних величин за результатами їх непрямих вимірювань. Патент 100288 Україна, МПК G06F 17/00, G06F 17/16 № а 2011 01844 заявл. 17.02.2011; опубл. 10.12.2012, Бюл. №23.

463. Згуровский М.З., Бидюк П.И., Терентьев А.Н. Методы построения байесовских сетей на основе оценочных функций. *Кибернетика и системный анализ*. 2008. N 2. С. 81–88.

464. Куссуль Э.М. Ассоциативные нейроразподобные структуры. Киев: Наукова думка, 1992. 144 с.

465. Рачковський Д.А., Гриценко В.І. Спосіб та система перетворення векторних даних. Патент 102980 Україна, МПК G06F 17/14. № а 2012 12937 заявл. 14.11.2012; опубл. 27.08.2013, Бюл. №16.

466. Рачковський Д.А., Гриценко В.І., Місуно І.С., Сліпченко С.В. Спосіб перетворення масиву векторних даних і комп'ютерна система для його реалізації. Патент 104674 Україна, МПК G06F 17/14. № а 2012 12692 заявл. 04.11.2012; опубл. 25.02.2014, Бюл. №4.

467. Рачковский Д.А., Слипченко С.В., Мисуно И.С. Ядра Мерсера для кодирования числовых векторов случайными гиперпрямоугольными рецептивными полями в задачах классификации. *Proc. ICNC'09*. 2009. Vol. 2. P. 150–153.

468. Рачковский Д.А., Слипченко С.В., Мисуно И.С. Интеллектуальная обработка данных протеомики для прогнозирования чувствительности глиомы к химиотерапии. *Кибернетика и вычислительная техника*. 2010. Вып. 161. С. 90–105.

Перелік умовних позначень і скорочень

- $1\{\cdot\}$ — індикаторна функція
- $|\mathbf{A}|$ — кількість ненульових компонентів в бінарному векторі \mathbf{A}
- $\langle \mathbf{a}, \mathbf{b} \rangle$ — скалярний добуток векторів \mathbf{a}, \mathbf{b}
- D — розмірність вхідного вектора
- d — розмірність вихідного (перетвореного) вектора
- C_a^b — число поєднань з a по b
- CRS — conditional random sampling (умовне випадкове семплювання)
- dist — відстань
- dist_{Bhat} — відстань Бхаттачарія
- dist_{Cheb} — відстань Чебишева
- dist_{Euc} — відстань Евкліда
- dist_{Hel} — відстань Хеллінгера
- dist_{JS} — відстань Йенсена-Шеннона
- dist_{KL} — відстань Кульбака-Лейблера
- dist_{Mah} — відстань Махаланобіса
- dist_{Man} — манхеттенська відстань
- dist_{Mink^s} — відстань Мінковського порядку s
- dist _{χ^2} — відстань χ^2 -квадрат
- DJL — distributional Johnson-Lindenstrauss
- $E\{\cdot\}$ — математичне сподівання випадкової величини
- i.i.d. — independent and identically distributed (незалежно і однаково розподілена)
- JL — Johnson-Lindenstrauss (Джонсон-Лінденштраус)
- KLSH — kernel LSH (ядерний LSH)
- KPCA — kernel PCA (ядерний PCA)
- LS — leverage scores
- LSH — locality-sensitive hashing (локально-чутливе хешування)
- LSF — locality-sensitive filtering (локально-чутлива фільтрація)

- L_s — векторний простір з s -нормою $\|\mathbf{a}\|_s$ і відстанню $\text{dist}_{\text{Mink}}^s$
 M — кількість ненульових компонентів в бінарному векторі
 MLE — maximum likelihood estimation (оцінювання максимальної правдоподібності)
 N — кількість об'єктів
 $\text{nnz}(\mathbf{a})$ — кількість ненульових компонентів у векторі \mathbf{a}
 Norm — розподіл Гауса
 p — ймовірність ненульового компонента в бінарному векторі
 p_{equ} — ймовірність збігу компонентів бінарних векторів
 p_{cond} — умовна ймовірність збігу ненульових компонентів бінарних векторів
 p_{join} — ймовірність збігу ненульових компонентів бінарних векторів
 PCA — principal component analysis (метод головних компонент)
 $\text{Pr}\{\cdot\}$ — ймовірність події
 PSD — positive semi-definite (додатно напіввизначена)
 RFF — random Fourier features (випадкові ознаки Фур'є)
 RFM — random feature map (випадкове ознакове відображення)
 sim_{cos} — косинусна схожість (косинус кута між векторами)
 sim_{dot} — скалярний добуток
 sim_{Jac} — схожість Жаккара
 SVM — Support Vector Machine (метод опорних векторів)
 Unif — рівномірний розподіл
 $\nu(\mathbf{A}, \mathbf{B})$ — відносне перекриття бінарних векторів \mathbf{A}, \mathbf{B}
 $\text{Var}\{\cdot\}$ — дисперсія випадкової величини
 АПНМ — асоціативно-проективні нейронні мережі
 НС — найближчий сусід
 в.в. — випадкова величина
 ІС — індексна структура
 к.ф.р. — кумулятивна функція розподілу
 м.с. — математичне сподівання
 РП — розподілене подання
 с.к.о. — середньоквадратичне відхилення
 ЦГТ — центральна гранична теорема

Зміст

Передмова	3
Розділ 1. Основні поняття	7
1.1. Векторні подання об'єктів і міри їхньої відстані та схожості	9
1.1.1. Метричні простори і міри відстані та схожості	9
1.1.2. Вектори та векторні простори.....	10
1.1.3. Відстані між бінарними векторами	13
1.2. Формування початкового векторного подання даних	15
1.2.1. Подання початкових об'єктів у вигляді дійсних векторів.....	15
1.2.2. Подання початкових об'єктів у вигляді бінарних векторів.....	17
1.2.3. BigData — великі дані.....	18
1.3. Вкладення і скетчі та їхні типи	20
1.3.1. Вкладення.....	20
1.3.1.1. Точні та наближені вкладення	20
1.3.1.2. Спотворення вкладень	21
1.3.1.3. Типи вкладень.....	22
1.3.2. Скетчі.....	25
1.3.3. Взаємозв'язок скетчів і вкладень	27
1.3.3.1. Про еквівалентність скетчів та вкладень	28
1.4. Розподілені та локальні подання даних	30
Розділ 2. Дійсні вектори для швидкого оцінювання відстані та схожості	33
2.1. Зниження розмірності векторів евклідового простору за допомогою випадкових проєкцій	35
2.1.1. Леми Джонсона-Лінденштрауса (D)JL	35
2.1.1.1. Типи випадкового лінійного перетворення JL	37
2.1.1.2. Аналоги леми JL для скалярного добутку, кута і вкладень в манхеттенський простір.....	40
2.1.2. Прискорення випадкових проєкцій та дерандомізація	40
2.1.2.1. Випадкові розріджені матриці	41
2.1.2.2. Матричні конвеєри для швидкого перетворення JL.....	41
2.1.2.3. Оптимізоване розріджене перетворення JL.....	43
2.1.2.4. Дерандомізація випадкових проєкцій	45
2.1.3. Дисперсії оцінок евклідової відстані та скалярного добутку у разі проєкцій випадковими матрицями різного типу.....	47
2.1.3.1. Оцінювання скалярного добутку	49
2.1.3.2. Оцінювання евклідової відстані	51
2.1.3.3. Аналіз оцінок	53

2.1.3.4. Експериментальне дослідження	56
2.1.3.5. Висновки	58
2.2. Вкладення та скетчі для оцінювання неевклідових відстаней	
Мінковського	61
2.2.1. Про неможливість зниження розмірності для неевклідових відстаней	61
2.2.2. Вкладення для даних з обмеженнями та без зменшення розмірності	62
2.2.2.1. Вкладення для даних з обмеженнями	62
2.2.2.2. Вкладення всього векторного простору з відстанями Мінковського для $1 \leq s \leq 2$	63
2.2.2.3. Вкладення гаусовими випадковими проєкціями при $s > 2$	64
2.2.3. Скетчі для відстаней Мінковського $0 < s \leq 2$ на основі стійких випадкових проєкцій	65
2.2.3.1. Формування скетчів на основі стійких розподілів	65
2.2.3.2. Оцінювання евклідової відстані за вибірковим середнім і медіаною	66
2.2.3.3. Медіанне оцінювання для манхеттенської відстані та інших відстаней Мінковського	67
2.2.3.4. Підвищення точності оцінювання відстаней Мінковського для $0 < s \leq 2$	68
2.2.3.5. Аналоги леми JL для манхеттенської відстані	70
2.2.3.6. Розріджені стійкі випадкові проєкції	71
2.3. Ядерні схожості та їх апроксимація.....	73
2.3.1. Швидке оцінювання ядерних схожостей з адаптацією до даних.....	75
2.3.2. Швидке оцінювання ядерних схожостей за векторами без адаптації до даних	78
2.4. Інші дійсні вкладення та скетчі	82
2.4.1. Вкладення відстаней між розподілами	82
2.4.2. Вкладення для швидкого наближеного пошуку за схожістю.....	83
2.4.3. Зменшення розмірності для оцінювання евклідової відстані шляхом розбиття векторів на частини.....	84
2.5. Висновки	86
2.5.1. Переваги та недоліки вкладень векторів евклідового простору випадковою проєкцією	86
2.5.2. Переваги та недоліки скетчів, отриманих стійкими випадковими проєкціями.....	87
2.5.3. Порівняння зі скетчами, отриманими семплюванням.....	87
2.5.4. Методи оцінювання схожості з навчанням	88
Розділ 3. Бінарні вектори для швидкого оцінювання відстані та схожості.....	91
3.1. Бінаризовані вкладення для оцінювання кута між векторами.....	93
3.1.1. Спотворення оцінок кута за бінаризованими гаусовими випадковими проєкціями.....	95
3.1.2. Прискорення формування бінаризованих вкладень	96
3.1.2.1. Комбінація FJLE та гаусової матриці.....	97

3.1.2.2. Безпосереднє використання FJLE.....	98
3.1.2.3. Конвеєри зі згладжуванням вхідних векторів.....	98
3.1.2.4. Прискорення проєкції випадковими розрідженими матрицями.....	99
3.1.3. Формування бінаризованих вкладень великої розмірності.....	100
3.2. Формування бінарних скетчів за допомогою LSH і оцінювання ядерних схожостей.....	102
3.2.1. Бінарні та цілочисельні скетчі з LSH-функцій.....	102
3.2.2. Локально-чутливе хешування SimHash та KLSH.....	104
3.2.2.1. Сімейство LSH SimHash для кута між векторами.....	104
3.2.2.2. KLSH: LSH для ядерних схожостей на основі SimHash.....	104
3.2.2.3. Експерименти з KLSH.....	106
3.2.2.4. Hash-SVM як покращення KLSH.....	107
3.2.2.5. LSH для неметричних відстаней.....	107
3.2.3. Апроксимація ядерних схожостей за бінарними векторами.....	108
3.2.3.1. Бінаризовані вкладення для апроксимації інваріантних до зсуву ядер та адитивних гомогенних ядер.....	109
3.2.3.2. Формування бінарних векторів з використанням LSH для оцінювання ядерних схожостей.....	109
3.2.3.3. Наближені хеш-функції ядра.....	111
3.2.4. Сімейства LSH на основі стійких випадкових проєкцій.....	111
3.2.4.1. Скетчі для евклідової відстані.....	113
3.2.4.2. Бінарні скетчі на основі стійких випадкових проєкцій для інших відстаней.....	114
3.3. Інші типи бінарних та цілочисельних вкладень і скетчів.....	116
3.3.1. Вкладення із цілочисельним квантуванням.....	116
3.3.2. Скетчі для пошуку за схожістю.....	118
3.3.2.1. Бінарні скетчі для оцінювання манхеттенської відстані.....	118
3.3.2.2. Бінарні скетчі для оцінювання евклідової відстані.....	119
3.3.2.3. Асиметричні відстані між скетчами.....	119
3.3.2.4. Асиметричний скалярний добуток між скетчами.....	120
3.3.3. Компактне подання бази для оцінювання відстаней.....	122
3.3.4. Скетчі для оцінювання відстані Хемінга між щільними бінарними векторами.....	122
3.4. Висновки.....	124
3.4.1. Ефективність бінарних та цілочисельних векторів.....	124
3.4.2. Бінарні вектори у наближеному пошуку за схожістю.....	125
3.4.3. Адаптація до даних.....	126
Розділ 4. Кодвектори: рандомізовані розріджені бінарні векторні подання об'єктів.....	127
4.1. Кодвектори та їхні властивості.....	128
4.1.1. Формування кодвекторів і оцінювання схожості вхідних об'єктів.....	128
4.1.1.1. Формування кодвекторів за рандомізованими алгоритмами.....	129
4.1.1.2. Покомпонентне формування кодвекторів і оцінювання схожості за емпіричною ймовірністю збігу одиничних компонентів кодвекторів.....	132

4.1.1.3. Лінеаризація функції випадкового аргументу (дельта-метод) для визначення дисперсії оцінки схожості	135
4.1.1.4. Оцінки схожості за невідомої залежності ймовірності збігу одиничних компонентів кодвекторів від міри схожості вхідних подань.....	136
4.1.1.5. Оцінка схожості за оцінкою умовної ймовірності збігу одиничних компонентів кодвекторів	137
4.1.2. Типи методів формування кодвекторів.....	141
4.2. Методи формування кодвекторів числових даних на основі випадкових проєкцій.....	144
4.2.1. Формування кодвекторів гаусовою випадковою матрицею з пороговим перетворенням	147
4.2.2. Геометрична інтерпретація	151
4.2.3. Оцінювання мір схожості вхідних векторів за кодвекторами і помилки оцінок.....	152
4.2.3.1. Оцінки кута і помилки оцінок.....	152
4.2.3.2. Оцінки скалярного добутку і евклідової відстані та помилки оцінок ..	155
4.2.4. Проекція випадковими матрицями з дискретними елементами	155
4.2.4.1. Збіжність до гаусового розподілу.....	156
4.2.4.2. Швидкість збіжності до гаусового розподілу	160
4.2.4.3. Експериментальне дослідження швидкості збіжності до гаусового розподілу	162
4.2.4.4. Експериментальне дослідження помилки оцінок кута	165
4.2.4.5. Експериментальне дослідження оцінок мір схожості за кодвекторами	167
4.2.5. Варіанти реалізації формування кодвекторів на основі випадкової проєкції.....	172
4.2.5.1. Нейромережева реалізація.....	172
4.2.5.2. Комп'ютерна реалізація з випадковою матрицею.....	173
4.2.5.3. Реалізація без зберігання матриці.....	175
4.2.6. Висновки	176
Післямова.....	178
Список літератури	183
Перелік умовних позначень і скорочень	210
Зміст	212