

Cómo escribir artículos



Dificultades para escribir artículos

- Publicar artículos es una actividad crítica para los investigadores
- Por otro lado, publicar es una tarea difícil
 - Pocos artículos son aceptados
 - Se compite contra buenos artículos
 - Los revisores son potenciales competidores

¿Qué es un artículo?

- Un artículo o paper es un relato en donde se presenta la **motivación**, los **métodos**, los **resultados** y las **discusiones** de una investigación científica o de un desarrollo tecnológico.
- Estos cuatro elementos deben estar presentes en el **resumen**, en la **introducción** y en la **conclusión**.
 - Cada sección tiene su propia finalidad y presenta esta información con diferentes niveles de detalle.
 - No se trata de una repetición.
 - Sirve para que los investigadores puedan “leer” decenas, centenas o millares de artículo por año.
 - Cada una de estas partes contiene una versión integral del artículo.

Partes independientes de un artículo

- Las **versiones integrales** de un artículo, que pueden ser leídas independientemente sin perder la visión general, son:
 - Resumen
 - Introducción
 - Conclusión

Partes independientes de un artículo

- El **resumen**, **resume** el artículo, o sea presenta la **motivación** (contexto y definición del problema), el **método** (cómo fue realizado, con que materiales, con que métodos), los **resultados** (e solamente resultados sin hablar de los método y sin anticipar la discusión) y la **discusión** (es una interpretación de los resultados, una discusión sobre su impacto y significado).
- La **introducción**, extiende la contextualización y definición del problema, antes de pasar para la próxima sección. Conduce al lector a través de un párrafo de transición, en el cual se describe todo el contenido del artículo.
- La **conclusión**, también resume en artículo pero pone énfasis en las principales conclusiones y presenta recomendaciones para nuevos trabajos.

Partes independientes de un artículo

- Además de estas tres partes importantes, es de suma importancia que el **título** sea el mejor resumen del trabajo.
- El título debe ser concreto y preciso evitando las abreviaciones, solo usa aquellas de uso claramente extendido.
- Use adjetivos que describan las características de su trabajo
- Atribución, quién hizo el trabajo:
 - Nombres de los autores
 - Universidad, instituto u organización a la cual pertenecen
 - E-mail
 - Fecha

EJEMPLOS

Shankar, Krauth, Pu , Jonas ,Venkataraman, Stoica, Ragan-Kelley(2018)

numpywren: Serverless Linear Algebra

Background. We design numpywren to target linear algebra workloads that have execution patterns similar to Cholesky decomposition. Our goal is to adapt to the amount of parallelism when available and we approach this by decomposing programs into fine-grained execution units that can be run in parallel. To achieve this at scale in a stateless setting, we propose performing dependency analysis in a decentralized fashion. We distribute a global dependency graph describing the control flow of the program to every worker. Each worker then locally reasons about its downstream dependencies based on its current position in the global task graph.

Methods: LAMBDAPACK, our domain specific language for specifying parallel linear algebra algorithms. Classical algorithms for high performance linear algebra are difficult to map directly to a serverless environment as they rely heavily on peer-to-peer communication and exploit locality of data and computation – luxuries absent in a serverless computing cluster. Furthermore, most existing implementations of linear algebra algorithms like ScalPACK are explicitly designed for stateful HPC clusters. We thus design LAMBDAPACK to adapt ideas from recent advances in the numerical linear algebra community on expressing algorithms as directed acyclic graph (DAG) based computation.

Shankar, Krauth, Pu , Jonas ,Venkataraman, Stoica, Ragan-Kelley(2018)

numpywren: Serverless Linear Algebra

Results. For algorithms such as SVD and Cholesky that have variable parallelism, while our wall clock time is comparable (within a factor of 2), we find that numpywren uses 1.26x to 2.5x less resources. However for algorithms that have a fixed amount of parallelism such as GEMM, the excess communication performed by numpywren leads to a higher resource consumption. We configured three framework scaLAPACK, Dask and numpywren to minimize total resources consumed. We note that for ScaLAPACK and Dask this is often the minimum number of machines needed to fit the problem in memory. Compared to ScaLAPACK-512 we find that numpywren uses 20% to 33 % lower corehours.

Conclusions. We have presented numpywren, a distributed system for executing large-scale dense linear algebra programs via stateless function executions. We show that the serverless computing model can be used for computationally intensive programs with complex communication routines while providing ease-of-use and seamless fault tolerance, through analysis of the intermediate LambdaPACK language. Furthermore, the elasticity provided by serverless computing allows our system to dynamically adapt to the inherent parallelism of common linear algebra algorithms.

Phillips et al. (1991) Importance of the lay press in the transmission of medical knowledge to the scientific community. NEJM 325: 1180-1183.

Background. Efficient, undistorted communication of the results of medical research is important to physicians, the scientific community, and the public. Information that first appears in the scientific literature is frequently retransmitted in the popular press. Does popular coverage of medical research in turn amplify the effects of that research on the scientific community?.

Methods. To test the hypothesis that researchers are more likely to cite papers that have been publicized in the popular press, we compared the number of references in the Science Citation Index to articles in the New England Journal of Medicine that were covered by The New York Times with the number of references to similar articles that were not covered by the Times. We also performed the comparison during a three-month period when the Times was on strike but continued to prepare an “edition of record” that was not distributed; doing so enabled us to address the possibility that coverage in the Times was simply a marker of the most important articles, which would therefore be cited more frequently, even without coverage in the popular press.

Phillips et al. (1991) Importance of the lay press in the transmission of medical knowledge to the scientific community. NEJM 325: 1180-1183.

Results. Articles in the Journal that were covered by the Times received a disproportionate number of scientific citations in each of the 10 years after the Journal articles appeared. The effect was strongest in the first year after publication, when Journal articles received 72.8 percent more scientific citations than control articles. This effect was not present for articles published during the strike; articles covered by the Times during this period were no more likely to be cited than those not covered.

Conclusions. Coverage of medical research in the popular press amplifies the transmission of medical information from the scientific literature to the research community.

Esta división en 4 partes es común en artículos del área médica, pero aun cuando esta división no se encuentre de forma explícita debe ser posible observar esta estructura en los resúmenes.

Ejemplos de un resumen

- Phillips et al. comprobaron que la divulgación de la investigación médica en la prensa legla aumenta el impacto (número de citas) de las publicaciones científicas
 - ¿Cómo probaron eso?
 - Phillips et al. (1991) Importance of the lay press in the transmission of medical knowledge to the scientific community. NEJM 325: 1180-1183.

El Resumen

- El propósito del resumen es contarle a un lector potencial, en el menor espacio posible, que encontrará en el artículo.
- El resumen debe incluir:
 - Exposición clara de los objetivos
 - Descripción breve de los métodos
 - Resultados principales y su interpretación
- No debe incluir:
 - Referencias bibliográficas o a elementos dentro del artículo
 - Motivación o justificación
 - Detalles irrelevantes

¿Cómo debe ser escrito un artículo?

- Un artículo cuyo resumen no resuma, cuya introducción no introduzca al lector al tema, cuya conclusión no concluya, **no será leído** por investigadores calificados que reconocerán la falta de organización.
- En el texto científico la innovación debe estar en el **contenido** y no en la forma.
- **El lector no debe usar la imaginación**, este debe ser guiado a través del texto, para que pueda comprender a partir de la lectura, sin suspenso y sin confusión.

¿Cómo se buscan artículos?

- El potencial lector busca usando palabras claves generalmente en bibliotecas digitales y online.
- El potencial lector se ayuda en los títulos para seleccionar artículos entre decenas, centenas o quizás millares de documentos.
- Los títulos promisorios requerirán de la lectura del resumen. Tal vez se busque en la introducción una definición del problema con más detalle. Tal vez en algunos casos se quiera conocer las conclusiones.
- Es común que apenas una ínfima parte de los artículos consultados sean leídos completamente.
- Si el artículo está bien escrito, tendrá chances de ser leído, citado y recordado.

¿Qué califica a un artículo como bueno?

- Lo que califica a un artículo como bueno es el mérito de la **contribución** que ofrece a un área del conocimiento, o sea cuando los resultados representan una avance en el área.
 - Ver The illustrated guide to a Ph.D., de Matt Might, en <http://matt.might.net/articles/phd-school-in-pictures/>
- Relatar un descubrimiento relevante es la contribución de un **artículo de investigación**, pero ni toda contribución requiere ser novedosa. Según Smith (1990) hay otros dos tipos de contribución.
 - Estudio exhaustivo (survey paper)
 - Tutorial

¿Algunos principios universales a considerar para escribir un artículo?

- Algunos principios que deben considerarse:
 - Mérito del artículo (calidad del artículo - contenido)
 - Repetibilidad
 - Fundamentación
 - Legibilidad
 - Conciso
 - Objetividad
 - Clareza

¿Por donde empezar y como?

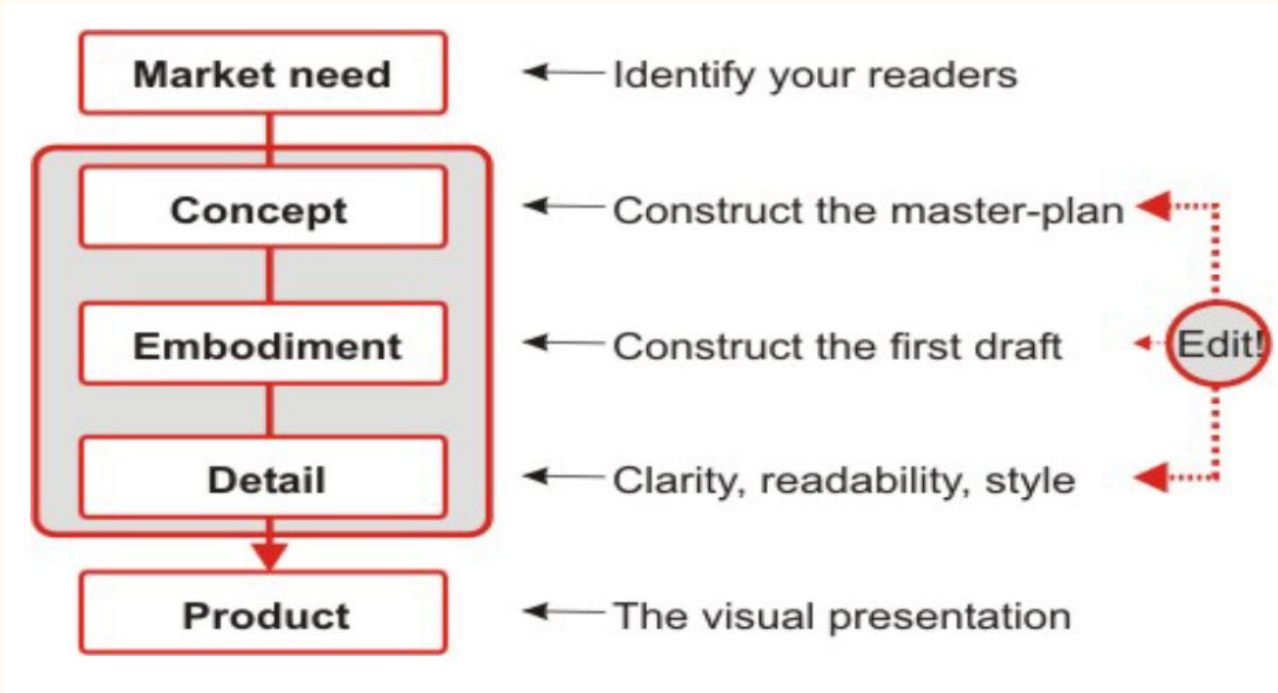
- Es difícil escribir un artículo cuando no se sabe lo que se quiere escribir.
- Para escribir un artículo de investigación, también es necesario que exista una investigación para comunicar.
- Las primeras actividades tratan sobre levantamiento de material suficiente para soportar el trabajo y planear su elaboración.

¿Por donde empezar y como?

- Típicamente el proceso para elaborar un artículo sigue los siguientes pasos:
 - Levantamiento bibliográfico
 - Mapa conceptual
 - Resumen preliminar
 - Primer borrador
 - Detallamiento
 - Validación
 - Reelaboración y sumisión y/o presentación
- **Antes de comenzar a escribir es necesario hacer cronograma.**

¿Por donde empezar y como?

- Diseño de un artículo según Ashby (2005)



¿Por donde empezar y como?

- **Levantamiento bibliográfico:** recolección de fuentes para soportar el artículo, para no correr el riesgo de presentar como novedad algo ya conocido por el lector. Es necesario revisar literatura especializada.
- **Mapa conceptual:** planeamiento del artículo. Puede ser hecho en una hoja A3/A4, de preferencia a mano. Puede ser diseñado juntamente con el resumen preliminar.
- **Resumen preliminar:** redacción de un texto de 100 palabras o menos. De una hasta tres oraciones completas (sujeto, verbo, complemento) como máximo para cada parte del resumen (motivación, métodos, resultados y discusión).

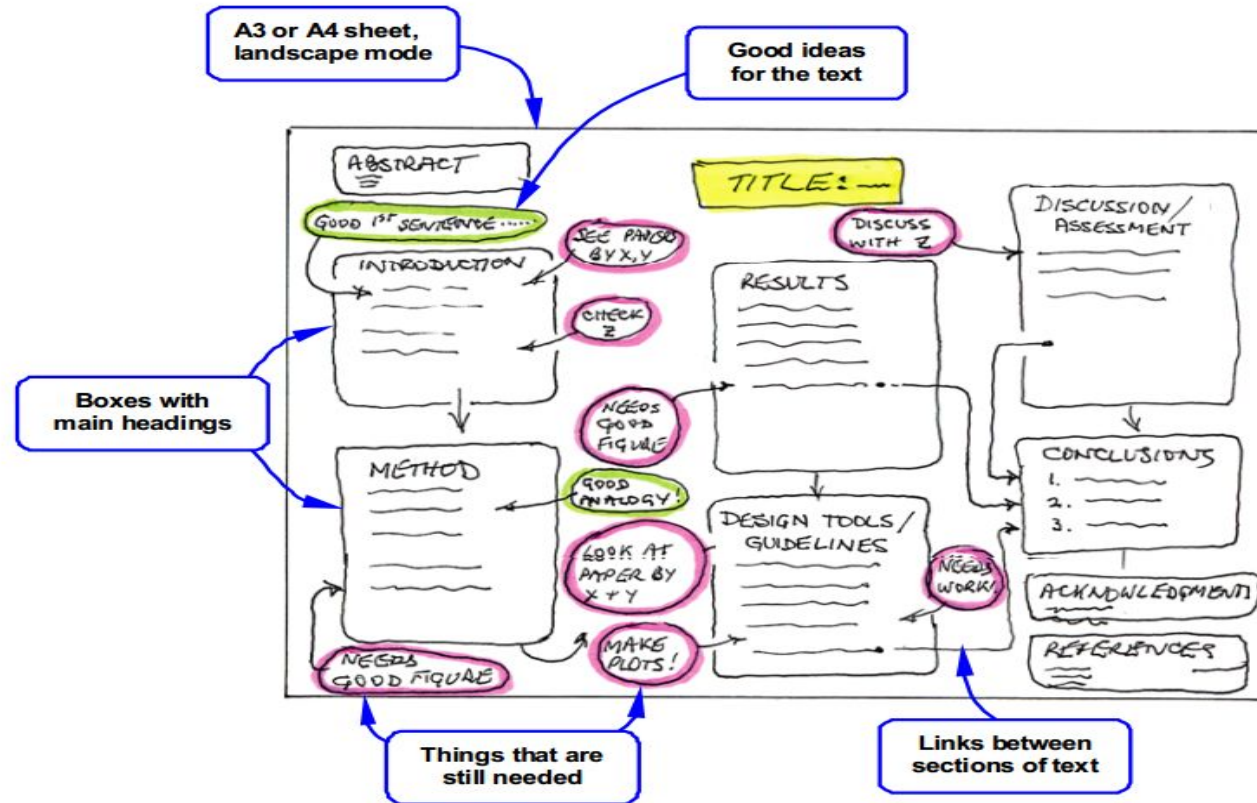
¿Por donde empezar y como?

- **Primer borrador:** elaboración de la primera versión completa del artículo. De preferencia en la forma de sección, figuras y ejemplos en lugar de texto corrido. Puede ser construido en la forma de slides.
- **Detallamiento:** mejorar la redacción para submeter el artículo a evaluación.
- **Validación:** feedback – revisión por pares, asesor, profesores.
- **Reelaboración y/o presentación:** apropiación del feedback, preparación del material para divulgación.

Mapa conceptual en planificación de artículos

- Ashby (2000) sugiere el uso de un mapa conceptual para **planificar** el artículo
 - Contiene la idea inicial de lo que se piensa escribir
 - Incluye las posibles secciones
 - Cada sección con sus probables ideas, figuras y ejemplos
- El objetivo del mapa conceptual es formar una **visión global**, para saber a dónde se va.
- Se puede **diseñar a mano**, reformulando hasta que parezca lo suficientemente maduro para que se pueda comenzar a detallar el artículo.

Mapa conceptual genérico de Ashby (2005)



¿Cuales son las partes típicas de un artículo

- **Introducción**
 - Identifica el problema.
 - Responde a las preguntas ¿qué? y ¿por qué?
- **Revisión de la Literatura**
 - Identifica lo que ya se conoce y lo que no se conoce
- **Métodos**
 - Responde a las preguntas ¿quién?, ¿cuándo? y ¿como?
- **Resultados**
 - Lo hallado sin calificación
 - Presentación “dura” de los datos (estadísticas, tablas, figuras)
- **Discusión**
 - Identifica lo que se ha encontrado
 - Describe el significado y las implicancias de los resultados
 - Acá no se habla sobre la revisión de la literatura pero sí se indica cómo estos hallazgos se encargan en el dominio revisado
- **Conclusión**

Introducción

- Cuál es el problema y por qué es interesante
- Cuáles han sido las principales contribuciones previas con las respectivas referencias (se debe poner en una sección a parte después de la introducción o antes de las conclusiones)
- Cuál es el aporte del artículo
- Una descripción breve del resto del artículo sección por sección.

Métodos

- Presenta métodos o técnicas propuestos (varía con el tipo de artículo como se ha mostrado en los ejemplos):
 - Presentación de algoritmos,
 - Presentación de una arquitectura (de hardware o de software),
 - Presentación de una metodología.
 - Resolución o demostración de teoremas, etc.
- Explica que es novedoso en su propuesta.
- Proporciona detalles suficientes de manera que el lector pueda reproducir lo que se hizo.
- No juntes los métodos con los resultados.

Resultados

- Describa el método experimental (si aplica al trabajo que se está realizando)
- Presenta los resultados de los experimentos, implementación de la arquitectura o aplicación de la metodología.
- Usa tablas y figuras, explicándolas de manera precisa en el texto.

Discusión

- Resultado principal
 - Énfasis en la discusión, resumir el resultado principal, ir de los datos numéricos a los datos descriptivos. En esta parte no se incluyen nuevos resultados.
- Interpretación de los resultados
 - ¿Qué significan los resultados?
- Interpretación en el contexto de la revisión de la literatura
 - ¿Son los resultados consistentes con investigaciones previas o con hallazgos previos?
- Implicaciones
 - ¿Cómo los resultados se pueden generalizar?
- Limitaciones
 - Adelantarse a las críticas de los revisores
- Resumen
 - Es la conclusión

Conclusiones

- Sucintamente resumir las implicaciones de los resultados como se mencionó anteriormente.
- No hacer declaraciones generales o conclusiones que van más allá de sus datos
- Presentar el valor del estudio
- Especificarle al lector lo que se debe llevar del artículo.

Acerca de las referencias

- Cita trabajos previos significativos
- Cita las fuentes de teorías, datos, técnicas y cualquier otra cosa que hayas tomado de otro lado
- Las referencias deben ser completas, siga un estándar específico:
 - APA: <http://www.citationmachine.net/apa/cite-a-book>
 - IEEE: <https://www.ieee.org/conferences/publishing/templates.html>
 - ACM: <https://www.acm.org/publications/proceedings-template>

Acerca de los apéndices

- Material esencial que interrumpiría el flujo del texto principal.
- Ejemplos:
 - Derivaciones matemáticas largas o teoremas a utilizar.
 - Tablas de datos
 - Descripción de procedimientos
 - Algoritmos

Consideraciones de estilo

- Sean claros y concisos
 - Digan lo que quieren decir de manera clara y evitando adornarse innecesariamente. Sea breve y vaya directamente al punto.
- **El uso de palabras precisas es clave.** La terminología técnica tiene significados muy concreto, aprenda cómo usarla de manera apropiada.
- Hoy en día es válido usar primera persona, si embargo, limita su uso. Generalmente es más apropiado en la introducción y en la discusión.
- Defina todos los símbolos y abreviaturas.
- Nunca te disculpes, sugiere mala planeación o incompetencia.

El trabajo final

- La apariencia visual es importante:
 - Buena organización
 - Encabezados claros
 - Figuras bien diseñadas
 - Tablas claras
- Verifique que el artículo se adapte al formato requerido (tipos de letra, numeración de secciones, número de columnas, estilo de referencias, ubicación de figuras y tablas, etc.)
- Considere el uso de un editor que facilite seguir el formato.

¿Qué tipo de literatura existe?

- **Literatura primaria:** primer registro publicado de la investigación. Puede ser artículo publicado en una revista, evento, tesis/disertación, etc.
- **Fuentes secundarias:** son síntesis, compilaciones y otros trabajos derivados de las fuentes primarias. Por ejemplo bibliografías, surveys, artículos de revisión, etc.
- **Fuentes terciarias:** no contiene contribuciones originales ni tampoco se refiere a ellas indirectamente, simplemente las discute. Por ejemplos los libros.
- **Literatura gris:** es investigación “no publicada”, que “no aparece” (en los principales indexadores y bases de datos). A veces puede incluir tesis, revistas, white papers, preprints, memoria de eventos etc.

¿Qué tipo de literatura es?

Red bayesiana

Una **red bayesiana**, **red de Bayes**, **red de creencia**, **modelo bayesiano (de Bayes)** o **modelo probabilístico en un grafo acíclico dirigido** es un **modelo grafo probabilístico** (un tipo de **modelo estático**) que representa un conjunto de **variables aleatorias** y sus **dependencias condicionales** a través de un **grafo acíclico dirigido** (DAG por sus siglas en inglés). Por ejemplo, una red bayesiana puede representar las relaciones probabilísticas entre enfermedades y síntomas. Dados los síntomas, la red puede ser usada para computar la probabilidad de la presencia de varias enfermedades.

Formalmente, las redes bayesianas son **grafos dirigidos acíclicos** cuyos nodos representan **variables aleatorias** en el sentido de **Bayes**: las mismas pueden ser cantidades observables, **variables latentes**, parámetros desconocidos o hipótesis. Las aristas representan dependencias condicionales; los nodos que no se encuentran conectados representan variables las cuales son condicionalmente independientes de las otras. Cada nodo tiene asociado una **función de probabilidad** que toma como entrada un conjunto particular de valores de las variables **padres** del nodo y devuelve la probabilidad de la variable representada por el nodo. Por ejemplo, si por padres son m variables booleanas entonces la función de probabilidad puede ser representada por una tabla de 2^m entradas, una entrada para cada una de las 2^m posibles combinaciones de los padres siendo verdadero o falso. Ideas similares pueden ser aplicadas a grafos no dirigidos, y posiblemente cíclicos; como son las llamadas **redes de Markov**.

Existen algoritmos eficientes que llevan a cabo la **inferencia** y el **aprendizaje** en redes bayesianas. Las redes bayesianas que modelan secuencias de variables (ej **señales del habla** o **secuencias de proteínas**) son llamadas **redes bayesianas dinámicas**. Las generalizaciones de las redes bayesianas que pueden representar y resolver problemas de decisión bajo incertidumbre son llamados **diagramas de influencia**.

¿Qué tipo de literatura es?

3.3. Model evaluation: quantifying the quality of predictions

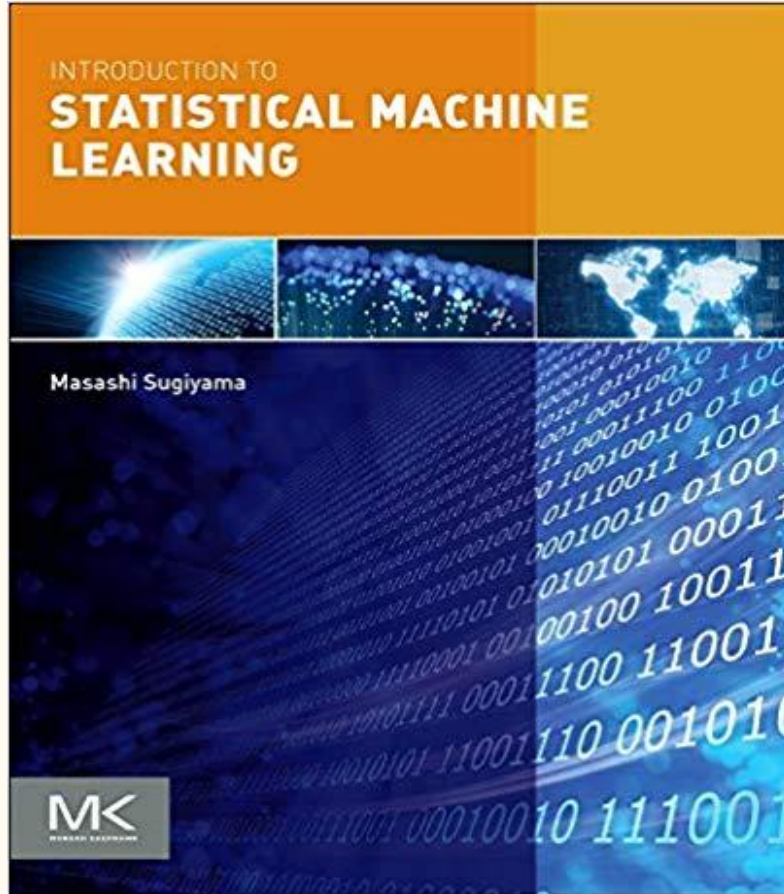
There are 3 different APIs for evaluating the quality of a model's predictions:

- **Estimator score method:** Estimators have a `score` method providing a default evaluation criterion for the problem they are designed to solve. This is not discussed on this page, but in each estimator's documentation.
- **Scoring parameter:** Model-evaluation tools using [cross-validation](#) (such as `model_selection.cross_val_score` and `model_selection.GridSearchCV`) rely on an internal *scoring* strategy. This is discussed in the section [The scoring parameter: defining model evaluation rules](#).
- **Metric functions:** The `metrics` module implements functions assessing prediction error for specific purposes. These metrics are detailed in sections on [Classification metrics](#), [Multilabel ranking metrics](#), [Regression metrics](#) and [Clustering metrics](#).

Finally, [Dummy estimators](#) are useful to get a baseline value of those metrics for random predictions.

See also: For "pairwise" metrics, between *samples* and not estimators or predictions, see the [Pairwise metrics, Affinities and Kernels](#) section.

¿Qué tipo de literatura existe?



Introduction to Statistical Machine Learning
Morgan Kaufmann; 1 edition (October 9, 2015)
Masashi Sugiyama.

¿Qué tipo de literatura existe?

arXiv.org > stat > stat.ML

Search or Article ID

All fields

[Help](#) | [Advanced search](#)

Machine Learning

Authors and titles for recent submissions

- [Wed, 24 Oct 2018](#)
- [Tue, 23 Oct 2018](#)
- [Mon, 22 Oct 2018](#)
- [Fri, 19 Oct 2018](#)
- [Thu, 18 Oct 2018](#)

[total of 206 entries: [1-25](#) | [26-50](#) | [51-75](#) | [76-100](#) | ... | [201-206](#)]
[showing 25 entries per page: [fewer](#) | [more](#) | [all](#)]

Wed, 24 Oct 2018 (showing first 25 of 38 entries)

[1] [arXiv:1810.09920](#) [[pdf](#), [other](#)]

Clustering Time Series with Nonlinear Dynamics: A Bayesian Non-Parametric and Particle-Based Approach
[Alexander Lin](#), [Yingzhuo Zhang](#), [Jeremy Heng](#), [Stephen A. Allsop](#), [Kay M. Tye](#), [Pierre Jacob](#), [Demba Ba](#)
Comments: main text, supplementary material
Subjects: **Machine Learning** ([stat.ML](#)); Machine Learning ([cs.LG](#)); Neurons and Cognition ([q-bio.NC](#)); Computation ([stat.CO](#))

[2] [arXiv:1810.09912](#) [[pdf](#), [other](#)]

Efficient Bayesian Experimental Design for Implicit Models
[Steven Kleinmenges](#), [Michael Gutmann](#)
Subjects: **Machine Learning** ([stat.ML](#)); Machine Learning ([cs.LG](#)); Computation ([stat.CO](#)); Methodology ([stat.ME](#))

[3] [arXiv:1810.09899](#) [[pdf](#), [other](#)]

Dynamic Likelihood-free Inference via Ratio Estimation (DIRE)
[Traiko Dinev](#), [Michael U. Gutmann](#)
Comments: For a demo, see [this https URL](#)
Subjects: **Machine Learning** ([stat.ML](#)); Machine Learning ([cs.LG](#)); Computation ([stat.CO](#)); Methodology ([stat.ME](#))

¿Donde hacer la búsqueda bibliográfica?

- ScienceDirect: <http://www.sciencedirect.com/>
- ACM DL: <http://dl.acm.org/>
- IEEExplorer:
<http://ieeexplore.ieee.org/Xplore/guesthome.jsp?reload=true>
- Scopus: <http://www.scopus.com/home.url>
- Google Scholar: <http://scholar.google.com>
- ISI Web of Science
- ArXiv: <https://arxiv.org/>

Referencias

- Ashby, M. (2000) How to write a paper. Engineering Department, University of Cambridge, Version 6.
- María de los Ángeles Fernández, Julio del Valle. Cómo iniciarse en la investigación académica. Una guía práctica. 2017. PUCP.