به نام خدا

کاربردها و چالش های تحلیل کلان داده ها در شهر هوشمند

مولفان :

زهرا گلشنی

عارف قائدی نیا

انتشارات ارسطو

(سازمان چاپ و نشر ایران – ۱۴۰۲)

نسخه الکترونیکی این اثر در سایت سازمان چاپ و نشر ایران و اپلیکیشن کتاب رسان موجود می باشد

chaponashr.ir

سرشناسه : گلشنی، زهرا، ۱۳۵۸
عنوان و نام پدیدآور : کاربردها و چالش های تحلیل کلان داده ها در شهر هوشمند/ مولفان زهرا گلشنی، عارف قائدی نیا.
مشخصات نشر : ارسطو (سازمان چاپ و نشر ایران)، ۱۴۰۲.
مشخصات ظاهری : ۴۷ ص.
شابک : ۱-۹۰۹-۳۳۹-۶۲۲-۹۷۸
وضعیت فهرست نویسی : فیپا
یادداشت : کتابنامه : ص ۴۷ .
موضوع : شهر هوشمند - تحلیل کلان داده ها - کاربردها و چالش ها
شناسه افزوده : قائدی نیا، عارف، ۱۳۵۵
رده بندی کنگره : LC۴۷۱۳
رده بندی دیویی : ۳۷۱/۹۰۷۱۳
شماره کتابشناسی ملی : ۹۴۲۳۱۱۳
اطلاعات رکورد کتابشناسی : فیپا

نام کتاب : کاربردها و چالش های تحلیل کلان داده ها در شهر هوشمند
مولفان : زهرا گلشنی - عارف قائدی نیا
ناشر : ارسطو (سازمان چاپ و نشر ایران)
صفحه آرایی، تنظیم و طرح جلد: پروانه مهاجر
تیراژ : ۱۰۰۰ جلد
نوبت چاپ : اول – ۱۴۰۲
چاپ : زبرجد
قیمت : ۴۷۰۰۰ تومان
فروش نسخه الکترونیکی - کتاب‌رسان :
https://chaponashr.ir/ketabresan
شابک : ۱-۹۰۹-۳۳۹-۶۲۲-۹۷۸
تلفن مرکز پخش : ۰۹۱۲۰۲۳۹۲۵۵
www. chaponashr. ir

انتشارات ارسطو
چاپ و نشر ایران
Chaponashr.ir

فهرست مطالب

عنوان **صفحه**

تحلیل کلان داده ها در شهر هوشمند ۵
اهداف و کاربردها ۶
آموزش هوشمند ۶
چراغ‌های ترافیک هوشمند ۷
شبکه‌ی هوشمند ۸
مسایل و مشکلات ۱۰
منابع داده و ویژگی‌هایش ۱۱
اشتراک گزاری اطلاعات و داده ۱۱
امنیت و حریم خصوصی ۱۲
هزینه ۱۲
جمعیت شهر هوشمند ۱۳
راه حل ها ۱۳
تحلیل کلان داده های مبتنی بر ابر برای آینده ی شهر هوشمند ۱۳
یک طراحی معماری انتزاعی از تجزیه‌وتحلیل‌های کلان داده‌ی مبتنی بر ابر ۱۵
یک مورد کاربرد تحلیلی با استفاده از داده‌ی باز Bristol ۱۹
پیاده‌سازی نمونه‌ی اولیه ۲۱
راه‌اندازی آزمایشی و نتایج ۲۴
یک چارچوب تحلیل کلان داده‌های جدید برای خانه‌های هوشمند ۲۵
پیاده‌سازی نمونه‌ی اولیه مبنی بر Hadoop ۲۹
لایه‌ی ادغام داده ۳۱
لایه‌ی تحلیل داده ۳۲
لایه‌ی کاربرد هوشمند ۳۳
لایه‌ی امنیتی ۳۳
اصول طراحی SCDAP ۳۴
مقایسه با رویکردهای سنتی ۳۵
برنامه‌ریزی شهری و ساخت شهر هوشمند بر اساس اینترنت اشیا و تحلیل‌های کلان داده ۳۸

مدل پیاده‌سازی و معماری تحلیلی کلان داده ۳۹
نتیجه گیری ۴۳
منابع ۴۷

تحلیل کلان داده ها در شهر هوشمند

اصطلاح کلان داده‌ها برای توصیف رشد، تکثیر، ناهمگونی، پیچیدگی، دسترسی، زمان‌بندی، تغییرپذیری و استفاده از داده‌ها در حوزه‌های کاربردی مختلف استفاده می‌شود .ویژگی‌های مذکور پردازش کلان داده‌ها را از قابلیت‌های محاسباتی و تحلیلی برنامه‌های کاربردی استاندارد و زیرساخت‌های مرسوم پایگاه داده خارج می‌کند. سیستم‌های تحلیلی سنتی برای مدیریت کلان داده‌ها مناسب نیستند. این بدان معنی است که پردازش کلان داده‌ها شامل استفاده از ابزار (طبقه‌بندی، خوشه‌بندی، رگرسیون و سایر الگوریتم‌ها)، تکنیک‌ها (داده‌کاوی، یادگیری ماشین، تجزیه‌وتحلیل آماری و...) و فناوری‌ها (هادوپ، اسپارک، اچبیس، مانگودی بی) است که فراتر از محدودهٔ روش‌های تحلیلی است که در استخراج دانش مفید از حجم وسیع داده‌ها برای تصمیم‌گیری دقیق و سریع باهدف افزایش بینش استفاده می‌شود. هیچ تعریف متمرکز یا قطعی از کلان داده‌ها درزمینهٔ شهرهای هوشمند پایدار وجود ندارد. این اصطلاح را می‌توان برای توصیف حجم بسیار زیاد از اطلاعات شهری به کار برد که دست‌کاری، تحلیل، مدیریت و ارتباط آن‌ها نیازمند هماهنگی و محاسبات تحلیلی و لجستیکی تدارکاتی بسیاری است .گفتنی است که چنین اطلاعاتی با برچسب‌های مکانی و زمانی روبه‌رو هستند که عمدتاً از انواع مختلف حس‌گرها و به‌طور خودکار تولید می‌شوند .به‌رغم عدم توافق در تعریف کلان داده‌ها، به نظر می‌رسد این توافق وجود دارد که کلان داده‌ها به پیشرفت‌های عظیم، نوآوری‌ها، امکانات فراوان و فرصت‌های شگفت‌انگیز در آینده منجر خواهند شد. ویژگی‌های اصلی کلان داده‌ها حجم بسیار زیاد داده، سرعتی که بتوان آن را پردازش کرد و طیف گسترده از انواع داده‌هاست. اصطلاح تحلیل کلان داده‌ها به‌طورکلی به هر مقدار گسترده از داده‌ها اشاره دارد که به‌صورت بالقوه برای جمع‌آوری، ذخیره، بازیابی، یکپارچه‌سازی، انتخاب، پیش‌پردازش، تبدیل، تجزیه‌وتحلیل و تفسیر برای کشف دانش جدید مفید استفاده می‌شود(Bibri, 2018).

درزمینهٔ شهرهای هوشمند پایدار، تجزیه‌وتحلیل کلان داده‌ها به مجموعه‌ای از برنامه‌های کاربردی نرم‌افزاری پیچیده و اختصاصی و سیستم‌های پایگاه داده اشاره می‌کند که توسط

ماشین‌آلات با قدرت پردازش بسیار زیاد اجرا می‌شود و می‌تواند کلان داده‌ها را به دانش مفید برای تصمیم‌گیری آگاهانه تبدیل کند(ایجاد و افزایش بینش دربارهٔ حوزه‌های مختلف شهری مانند حمل‌ونقل، تحرک، ترافیک، محیط‌زیست، انرژی، برنامه‌ریزی و طراحی.) انواع رایج پردازش کلان داده‌ها عبارت‌اند از تجزیه‌وتحلیل پیش‌بینی، تشخیصی، توصیفی و تجربی که برای استخراج انواع دانش یا بینش از مجموعهٔ کلان داده‌ها به کار می‌رود و برای اهداف مختلف، بسته به دامنهٔ کاربرد آن‌ها، استفاده می‌شود. تجزیه‌وتحلیل شهری شامل استفاده از تکنیک‌های مختلف مرتبط با داده ازجمله تفکر داده‌های تحلیلی و اصول استخراج دانش مفید با استفاده از الگوهای پنهان و همبستگی‌های معنی‌دار، یادگیری ماشین، داده‌کاوی، تجزیه‌وتحلیل آماری، تجزیه‌وتحلیل رگرسیون (مدل‌سازی توضیحی در مقایسه با مدل پیش‌بینی)، پرس‌وجوی پایگاه داده، ذخیره‌سازی داده‌ها و یا ترکیبی از این تکنیک‌هاست. استفاده از این تکنیک‌ها بستگی به حوزهٔ شهری و همچنین ماهیت مشکلی دارد که باید حل‌وفصل شود.

اهداف و کاربردها

مثال‌های زیادی از به‌کارگیری کلان داده‌ها در شهر هوشمند وجود دارد که برخی از آن‌ها عبارت‌اند از:

آموزش هوشمند

ICT راه‌حلی را برای بهبود اثربخشی، کارایی و بهره‌وری فرایندهای آموزشی با استفاده از سرویس‌های هوشمند آموزشی که انعطاف‌پذیر و هوشمند هستند فراهم می‌کند تا از اطلاعات بهتر استفاده کنند، کنترل و ارزیابی را بهبود ببخشند و پشتیبانی بهتری را برای یادگیری بلندمدت برای تمام انسان‌ها فراهم کنند. اپلیکیشن های آموزش هوشمند انسان‌ها را در محیط‌های یادگیری فعال درگیر می‌کند و امکان تغییرات سریع جامعه و محیط را برای آن‌ها فراهم می‌کند. علاوه بر این، با تکیه‌بر کلان داده‌های جمع‌آوری‌شده در این زمینه و پردازش درست برای تولید اطلاعات موردنیاز، تأثیر مثبتی بر روی سطوح دانش و ابزارهای یادگیری برای تحویل یا حصول دانش ایجاد می‌شود. علاوه بر این، فناوری می‌تواند چنین فرصت‌هایی را همه‌جا در دسترس قرار دهد ازجمله مناطق دورافتاده و

روستایی که در آن رفت‌وآمد به مدارس ممکن نیست و یا وضعیت اقتصادی مردم ضعیف است و نمی‌توانند به مدل‌های گران‌قیمت بپردازند(Athawale & Gothawal, 2017).
با استفاده از ICT و کلان داده می‌توان به ایجاد یک جامعه‌ی مبتنی بر دانش کمک کرد که در آن می‌توان توانایی رقابت‌پذیری کشور را افزایش داد. کلان داده در آموزش به‌طور عمده با جمع‌آوری داده‌های مربوط به افراد (مثلاً دانش آموزان، معلمان، والدین، مدیران و دیگر پرسنل پشتیبانی)، زیرساخت (مانند مدارس، کتابخانه‌ها، امکانات رایانه‌ای، مکان‌های آموزشی، موزه‌ها، دانشگاه‌ها و دیگر نهادهای مرتبط) و اطلاعات (به‌عنوان‌مثال دوره‌ها، کتاب‌ها، امتحانات، نمرات، نظرسنجی‌ها، ارزیابی‌ها، گزارش‌ها و غیره) تولید می‌شود. این داده‌ها می‌تواند یک منبع مفید را برای تجزیه‌وتحلیل و استخراج روندها و مدل‌های مفید ایجاد کند و از آن‌ها برای ارائه‌ی آموزش پیشرفته‌تر و بهتر استفاده کند. به‌عنوان‌مثال کلان داده‌ها، سازمان‌های آموزشی را برای شخصی‌سازی یادگیری پشتیبانی می‌کنند(Athawale & Gothawal, 2017).

چراغ‌های ترافیک هوشمند

یکی از جنبه‌های اصلی شهر هوشمند کنترل خوب جریان ترافیک در شهر است که سیستم‌های حمل‌ونقل را ارتقا می‌دهد و رفت‌وآمد شهروندان و الگوهای کلی ترافیک شهرها را بهبود می‌دهد. هنگامی‌که جمعیت افزایش پیدا می‌کند مسائل ترافیکی، آلودگی و مسائل اقتصادی رخ می‌دهد. به‌واسطه‌ی این، استفاده از چراغ‌های ترافیک هوشمند و سیگنال‌ها یکی از مهم‌ترین تکنیک‌هایی است که شهرهای هوشمند برای مقابله با حجم بالای ترافیک و ازدحام از آن استفاده می‌کنند. چراغ‌های ترافیک هوشمند و سیگنال‌ها باید در سراسر شبکه‌های ترافیکی متصل شوند تا اطلاعات بیشتری را در مورد الگوهای ترافیکی ارائه دهند. هر سنسور یک پارامتر متفاوت از جریان ترافیک را تشخیص می‌دهد (به‌عنوان‌مثال سرعت ماشین‌ها، تراکم ترافیک، زمان انتظار در چراغ‌ها و غیره). سیستم با توجه به مقادیر این پارامترها تصمیم می‌گیرد و دستورالعمل‌های مناسبی را برای چراغ‌ها و سیگنال‌ها صادر می‌کند. بنابراین هر چه اطلاعات بیشتری در دسترس این سیستم باشد

تصمیمات آگاهانه‌تری خواهد گرفت. درنتیجه، برای ارائه‌ی بهترین سرویس ممکن در چراغ‌های ترافیک هوشمند بهترین راه‌حل جمع‌آوری داده از چراغ‌های ترافیک در سراسر شهر و ایجاد سیستم‌های تصمیم‌گیری هوشمند با این داده‌هاست و این به استفاده از تحلیل‌های کلان داده‌ی زمان واقعی نیاز دارد(Lim, Kim, & Maglio, 2018).

شبکه‌ی هوشمند

شبکه‌ی هوشمند یک مؤلفه‌ی مهم از یک شهر هوشمند است. این‌یک سیستم شبکه‌ی الکترونیک بازسازی‌شده است که از فناوری اطلاعات و ارتباطات برای جمع‌آوری و عمل بر روی‌داده‌های موجود مانند اطلاعات در مورد رفتارهای تأمین‌کنندگان و مصرف‌کنندگان در یک شیوه‌ی خودکار استفاده می‌کند. این اثربخشی، قابلیت اطمینان، اقتصاد و پایداری تولید و توزیع قدرت برق را بهبود می‌بخشد. یک شبکه‌ی هوشمند از کنترل‌های از راه دور مبتنی بر کامپیوتر با فناوری ارتباطی دوطرفه میان تولیدکننده‌ی برق و مصرف‌کننده برای افزایش کارایی شبکه و قابلیت اطمینان از طریق سیستم خود نظارتی و بازخورد استفاده می‌کند. این شامل قرار دادن سنسورهای هوشمند و کنتورها در تولید، انتقال و توزیع سیستم‌ها به‌علاوه در نقاط دسترسی مصرف‌کنندگان برای دریافت داده‌های زمان واقعی در مورد تولید برق فعلی، مصرف و اشکالات است.

مدل‌های قیمت‌گذاری پویا را برای مصرف برق پیاده‌سازی می‌کند. این به قطع بالقوه‌ی برق ناشی از تقاضاهای بالای مصرف‌کنندگان کمک می‌کند. این می‌تواند اطلاعات نزدیک به زمان واقعی را به مصرف‌کنندگان در مورد مصرف انرژی آن‌ها ارائه دهد و به آن‌ها امکان مدیریت مصرف بر اساس نیازهایشان و مبلغ متناسب فراهم می‌کند. (Lim, Kim, & Maglio, 2018).

دستگاه‌های مصرف‌کنندگان مانند ماشین لباس‌شویی می‌توانند با کنترل خودکار آن‌ها در طول دوره‌های باقیمت کمتر بهتر عمل کنند. اگرچه شبکه‌ی هوشمند مزایای بالقوه‌ی زیادی دارد اما نیازمند جمع‌آوری حجم زیادی از داده از رویه‌ها، انتقالات، توزیع و مصرف برق است. علاوه بر این، نیاز به پردازش داده‌های جمع‌آوری‌شده دارد که تحلیل‌های کلان

داده در زمان واقعی برای ارسال برخی اطلاعات کنترلی برای بهبود عملکرد کلی سیستم قدرت الکترونیک در نظر گرفته‌شده است(Lim, Kim, & Maglio, 2018).

چندین نمونه از اپلیکیشن های کلان داده بررسی شدند که می‌توانند به‌عنوان راهنمایی برای توسعه‌ی اپلیکیشن های شهر هوشمند در نظر گرفته شوند. جدول ۲-۱ نشان می‌دهد که چگونه شهرها در سراسر دنیا از اپلیکیشن های کلان داده در مؤلفه‌های مختلف شهر هوشمند با پیاده‌سازی پروژه‌های شهر هوشمند واقعی استفاده می‌کنند. بررسی برخی از پیاده‌سازی‌های واقعی نشان می‌دهد که مزایایی در کلان داده وجود دارد که در مؤلفه‌های شهر هوشمند منعکس‌شده‌اند. جدول ۲-۲ این مزایا را در حوزه‌های کاربردی مختلف شهر هوشمند نشان می‌دهد

جدول۱- نمونه هایی از پروژه‌های کلان داده در مؤلفه‌های شهر هوشمند (Nuaimi, Neyadi, Mohamed, & Al-jaroodi, 2015)

Smart city components	Big Data Projects	Location
Transportation, Mobility, and Logistics	An accelerated-time simulation for traffic flow (ATISMART model) based on the use of smart traffic lights and signals as a part of a smart city project. Accelerated-time simulations for traffic flow should take into consideration three different factors: the city map, the cars, and the smart signals. To implement a smart traffic flow, there are some requirements to consider such as network sensors, traffic lights, and CAS as the mathematical core of the model and Java for the GUI [16].	-
Healthcare	"Ministry of Health and Welfare initiated the Social Welfare Integrated Management Network to analyze 385 different types of public data from 35 agencies and comprehensively manage welfare benefits and services provided by the central government, as well as by local governments, to deserving recipients" [23].	South Korea
Public safety	"The Ministry of Food, Agriculture, Forestry, and Fisheries and the Ministry of Public Administration and Security, or MOPAS, plan to launch the Preventing Foot and Mouth Disease Syndrome system, harnessing big data related to animal disease overseas, customs/immigration records, breeding farm surveys, livestock migration, and workers in the livestock industry" [23].	South Korea
	"In 2004, to address national security, infectious diseases, and other national concerns, the Singapore government launched the Risk Assessment and Horizon Scanning (RAHS) program within the National Security Coordination Centre. Collecting and analyzing large-scale data sets, it proactively manages national threats, including terrorist attacks, infectious diseases, and financial crisis. … A notable RAC application is exploration of possible scenarios involving importation of avian influenza into Singapore and assessment of the threat of outbreaks occurring throughout southeast Asia" [23].	Singapore
Education	NEdNet (National Education Network) is an integrated system including network infrastructure services, education information services (EIS), and learning services, which facilitate higher-order thinking skills, support learner-centered self-directed and tailored learning, and decision support [13].	Thailand
Natural resources & energy	The UK government established the Horizon Scanning Centre (HSC) in 2004 to improve the government's ability to deal with cross-departmental and multi-disciplinary challenges. In 2011, the HSC's Foresight International Dimensions of Climate Change effort addressed climate change and its effects on the availability of food and water, regional tensions, and international stability and security by performing in depth analysis on multiple data channels [23].	UK
Government & agency administration	"To manage real-time analysis of high volume streaming data, develop a massively scalable, clustered infrastructure. … For discovery and visualization of information from thousands of real-time sources, encompassing application development and systems management built on Hadoop, stream computing, and data warehousing" [12].	USA
	"In 2009, the U.S. government launched data.gov as a step toward government transparency and accountability. It is a warehouse containing 420,894 datasets covering transportation, economy, health care, education, and human services and the data source" [12].	
	"In 2011, Syracuse, NY, in collaboration with IBM, launched a Smarter City project to use big data to help predict and prevent vacant residential properties. Michigan's Department of Information Technology constructed a data warehouse to provide a single source of information" [12].	

جدول۲- مزایای کلان داده در مؤلفه‌های شهر هوشمند (Nuaimi, Neyadi, Mohamed, & Al-jaroodi, 2015)

Smart City Components	Benefits of Big Data in Smart City Components
Smart Healthcare	• Allow healthcare providers and practitioners to gather, analyze, and utilize patient information, which can also be used by insurance companies and some government agencies. • Support processing complex occurrences to monitor, analyze, and flag potential health issues either on a daily basis or on a demand basis. • Increase the amount and real-time nature of data gathered for certain patients' healthissues through smart devices, which are connected to the home or hospital to monitor attributes like blood pressure, blood sugar, and sleep patterns for accurate and timely responses to health issues and for a comprehensive patient history records.
Smart Energy	• Facilitate decision-making related to the supply levels of electricity in line with actual demand of the citizens and over all affecting conditions. • Allow forecasting in a near-real time manner through efficient analysis of the big data collected. • Align with strategic objectives (resource optimization) through specific pricing plans consistent with supplies, demand, and production models.
Smart Transportation	• Recognize traffic patterns by investigating real time data • Reduce main city roads' congestion by predicting traffic conditions and adjusting traffic controls. Through big data, the smart city will be able to reduce traffic and accidents by opening new roads, enhancing the infrastructure based on congestion data, and collecting information on car parking and alternative roads. • Reduce supply chain waste by associating deliveries and optimizing shipping movements. • Enable data streaming to process and communicate traffic information collected through sensors, smart traffic lights and on-vehicle devices to drivers via smartphones or other communication devices. • Big data can be used to send feedback for specific entities to take action to alleviate or resolve a traffic problem.
Smart Environment	• Provide weather information that will lead to improving the country's agriculture, better informing people of possible hazardous conditions, and better management of energy utilization by providing more accurate predictions on demand.
Smart Safety	• Provide detailed and spatial and temporal geographic area maps and help to easily determine whatever changes may happen. • Help predict future environmental changes or natural disasters like earthquake detection that will give an opportunity to save lives and resources.
Smart Education	• Optimize academic research; for instance, astronomer can now analyze a huge astronomy dataset using powerful computers instead of manual analyses. By analyzing and exploring high quality digital images taken from space, new discoveries may happen in the fields. This is applicable to many science and research fields such as medical experiments, manufacturing operations, environmental studies, and economic and financial analysis. • Behavior and matchmaking will lead to new knowledge. From assessment of graduates to online attitudes, each student generates a unique data track. By analyzing these data, education institutes can realize whether they are using their resources in the right places and producing the right results.
Smart Governance	• Support the integration and collaboration of different government agencies and combine or streamline their processes. This will result in more efficient operations, better handling of shared data, and stronger regulation management and enforcement. • Improve business decisions through big data analytics support. By researching a firm's behavior and economic growth in addition to its rivals and environment conditions, more appropriate and effective decisions related to employment, production, and location strategies can be made. • Publish new policies for the benefit of data owners (citizens) and producers (government agencies). Government agencies will help develop the quality of the data, while citizens will show how they can use the data and transfer it to new knowledge to enhance the quality of government services. • Help governments focus on the citizens' concerns related to health and social care, housing, education, policing, and other issues.

مسایل و مشکلات

چالش‌های زیادی برای طراحی، توسعه و استقرار اپلیکیشن‌های کلان داده برای شهر هوشمند وجود دارد. شهرهای هوشمند دارای محیط‌های بسیار پویا و تکاملی هستند؛ بنابراین، اجتناب یا حداقل کاهش ریسک‌ها در طراحی اپلیکیشن‌های هوشمند و توسعه‌ی شهرهای هوشمند ضروری است. همچنین برخی اختلاف‌نظرها در مورد تعاریف، استفاده و مزایای کلان داده برای شهرهای هوشمند وجود دارد. این مربوط به ابزارهای کلان داده،