



Surface Go

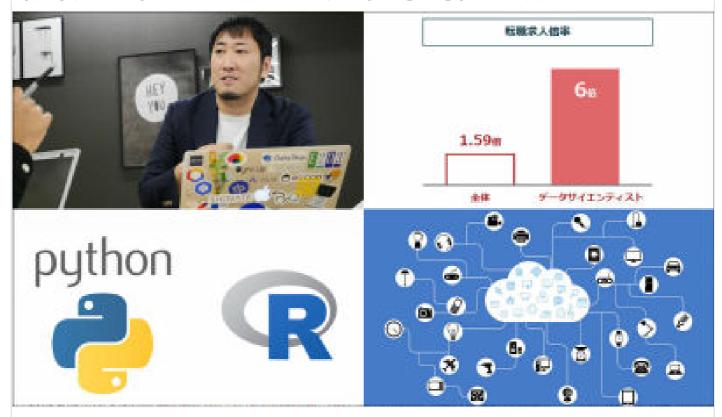


新しい Surface Go は、 Windows 10 (S モード)¹ ... ¥98,064

2018年11月21日 09時00分

広告

「むしろスキルがなければ給与が下がる可能性」、求められるデータサイエンティスト人材になる方法論をプロにイチからいろいろ聞いてみた



先端IT人材の中でも急増する需要に追い付かず、慢性的な「売り手市場」となっている「データサイエンティスト」という職種。企業から引く手あまたといわれるデータサイエンティスト職の実態はどうなっているのか?ということで、人材大手パーソルキャリアで現役のデータサイエンティストとして活躍する**鹿内学**さんに、いろいろと話を聞き、さらに世界最大級のオンライン学習教育プラットフォーム「Udemy(ユーデミー)」で役に立つ講座はどれかというのも教えてもらいました。

世界最大級のオンライン学習プラットフォーム | Udemy



目次

- **◆1:「データサイエンティスト」とは?**
- ◆2:「データサイエンティスト」が引く手あまたなワケ
- ◆3:データサイエンスが使えないと給料が減る時代が来る!?
- **◆4:「データサイエンティスト」になるための学習法とは?**
- ◆5:プロがおススメするUdemyデータサイエンス講座5選

研究者からビジネス界へ転身し、現在は人材大手パーソルキャリアの新規事業「Data Ship」というサービスの事業開発や転職サービス「**ミイダス**」のデータサイエンスプロジェクトを統括する鹿内学さん。



◆1:「データサイエンティスト」とは?

GIGAZINE(以下、「G」):

今回は、需要が急激に高まっているという「データサイエンティスト」という仕事について、データサイエンティストの学習支援、キャリア支援に携わっている視点から話を聞かせてください。まずは鹿内さんの簡単な自己紹介をお願いします。

鹿内学氏(以下、「鹿内」):

私は、奈良先端科学技術大学院大学で、理学の博士号を取得しました。当時は、ヒトの脳の研究をするために、眼球運動の計測やfMRIという脳活動計測を行っていました。最初の仕事では、京都大学の医学研究科で教鞭をとり、その後、情報学研究科にうつり、研究をしていました。研究者としての最後の2年は国のプロジェクトに関わっていて、「脳波を使ってドアとかを開けたり閉めたりする」というような研究をしていましたし、MRI、fMRI、NIRS、脳波(EEG)など様々な脳活動計測のデータ集積にも力を入れていました。

G:

「脳とコンピューターを繋ぐ」というような研究をされていたのですか?

鹿内:

そうです。ブレイン・コンピューター・インターフェース(BCI)とも言ったり、ブレイン・マシン・インタフェース(BMI)とも呼ばれる技術です。

G:

脳の研究から、データを扱うビジネスに転身されたのですか?

鹿内:

そうです。脳波だと「時系列データ」が出てくるのです。これを機械学習などで分析して、情報を取り出す。脳波のデータはすごくノイジーなので、データがあってもそこから情報を取り出すのは難しい。それを解決する技術や経験があり、ビジネスサイドで期待されたのが、データに関わることでした。

G:

アカデミックな世界からビジネスの世界に転身されるきっかけが何かあったのですか?

鹿内:

いくつかビジネスサイドからオファーを頂いたんです。

G:

もう既にその時点で「データサイエンティスト」として、という感じですか?

鹿内:

いいえ。当時は、自分自身がデータサイエンティストという職業だとは気づいてなかったですね。

G:

3、4年前はデータサイエンティストという言葉自体がそれほど一般的ではなかったですよね。

鹿内:

一般的ではないですね。まあ、言葉があることは知っていましたが、どういう職業なのかは詳しく知らないし、まさか自分がそれに該当するとは思ってもいなくて(笑)

G:

研究者の方で、自分がやっている研究が世間ではそういう風に評価されているとは知らない...... ということがあるのですね。

鹿内:

実は、今もそれはあって。博士課程の大学院生や研究者でも、自分ができるデータサイエンスの 仕事があると気づいていない人は多いです。だから「知る」ということ、「気づく」ということ は結構、重要だな、と感じています。



G:

データサイエンティストの"守備範囲"は、ものすごく広いのでしょうか?

鹿内:

広いですね。あと、作業の種類で定義できる職業でもないような気がしています。一般的な考えでないと思いますが、僕自身は、データがわかった上で事業開発をするのも、データサイエンティストの仕事だと考えています。

少し話しがそれるのですが、呼び方の問題もありますね。「データアナリスト」という言葉を使ったり、業界によっては「デジタルマーケター」と呼んだり、あとは「機械学習エンジニア」と呼んだり。言葉の使い方はまちまちで、私がアカデミックの世界にいたときは、「データサイエンティスト(笑)」という言い方もありました。このニュアンスはビジネスサイドでは通じなくて、ちょっとビックリしたんですけれど、データサイエンティストを揶揄するものです。「データサイエンティストってのは、この程度なのか」と。アカデミックな世界で機械学習の研究をやっている人からすると、圧倒的にレベルが低かったのですね。

僕自身も研究対象は認知神経科学なので、直接には機械学習の研究者ではありません。そこから見ても低いです。「データサイエンス」と呼んではいても、その仕事は、Excelで集計しているだけだったり、分析結果の解釈を明らかに間違えていたり……。極論すればグラフにしているだけ

の人たちもいた。それでデータサイエンティストと呼ばれる人たちがビジネスサイドでは、結構 目について、やっぱり言葉と実態が合わない状況が見えていましたね。そのせいで、所属部署に データサイエンスという名称でもついてない限り、データサイエンティストと自ら名乗る人があ まりいないのかもしれません。

G:

データサイエンティストという言葉自体を最近耳にするようになって、多くの人がイメージとしてはだいたい掴んでいるけれど、はっきりとは掴めてないというところがあると思います。もっとイメージしやすいようにデータサイエンティストが活躍することによってできたサービス、私たちが使っている具体的なサービスとは何ですか?

鹿内:

そうですね、例えばYahoo!ショッピングとかで買い物するときのリコメンド(オススメ)は、データサイエンスの成果ですね。最近では、AI内蔵の音声スピーカーとかありますよね。あれも機械学習を使っているはずです。それ以外にも、アビームコンサルティングさんなどコンサルティング(ファーム)でもデータサイエンスを使っています。また、人事的な情報に関してもどうやって異動・配置をしたり採用すれば良いのか、というのにデータサイエンスを使っています。

G:

基本的には機械学習に関わるものはすべてデータサイエンティストが関わっていると考えていい のですか?

鹿内:

データサイエンティストが関わっています。一般的な「線形回帰分析」という古典的な統計分析 法を使っても売り上げアップに貢献できるので、データサイエンスだからといって、必ず機械学 習を使うというわけではないのですが。

G:

簡単に言うと、情報の中から何か価値を生み出せる人が真のデータサイエンティスト、ということなのでしょうか?

鹿内:

そうです。データから価値を生み出すことは重要です。データサイエンティストというとよくデータの分析・評価のところに注目されがちなのですが、「データ活用設計」「データを取得・分析設計」から始まる場合もあります。



データ分析の業務フロー





Copyright 6: PERSON CARRES Co., Und. All Sights Reserved

30

鹿内:

例えばブリヂストンさんの場合、7、8年前にタイヤへセンサーをつけたのです。そもそもタイヤが高速で回転している中で、無線でデータを飛ばしながらデータを取ること自体が困難なのですが、データを取得できたおかげで、ビジネスができつつあります。タイヤのセンサーからのデータによって、路面が乾いているか、凍っているか、いろいろな状況が分かるのです。高速道路が凍結している場合には、融雪剤などを撒きますが、その判断をこれまではヒトが目視でやって、目視でも分からなければ路面を実際に人が触っていたそうです。けれど、高速道路でそういうことをすると危険じゃないですか。それがタイヤのセンサーから情報をとることで解決しているようです。

G:

なるほど。分析評価のところに視点が集まりがちだけれど、その前段階のデータを取るところから関わらないと最終的に良いものができないということですね。

鹿内:

良いものはできませんね。理想的には。



◆2:「データサイエンティスト」が引く手あまたなワケ

G:

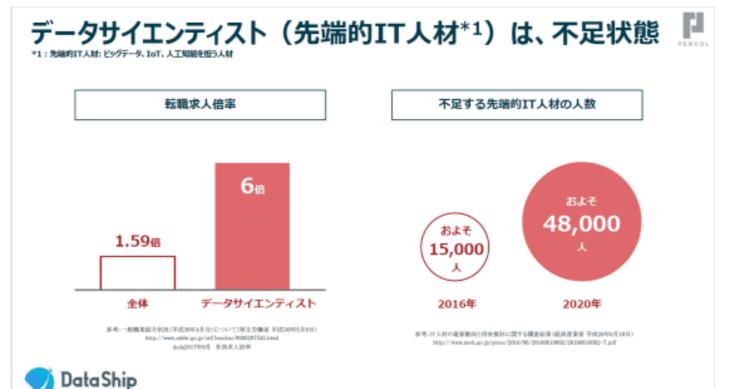
データサイエンティストの需要が急激に高まっているとのことですが、あらゆる分野で基本的に データをうまく活用することが求められているということなのですか?

鹿内:

求められています。それが「データサイエンティストが圧倒的に足りない」と言われていることに繋がっていると思います。

G:

「データサイエンティストの需要がものすごく高くて、転職求人倍率が6倍」というデータや、「先端IT人材が2020年には4万8000人も不足」という(PDFファイル)データがあります。現在、「先端IT人材」はどれくらいの数がいるのでしょうか?



鹿内:

経産省が出した2016年の**(PDFファイル)データ**では11万人くらいいます。それに対して、3万人くらいが不足しています。この「不足」が何を意味するかというと、「ここにデータサイエンティストがいれば、新しいビジネスができる」ということです。

Copyright © PERSOL CAREER Co., Ltd. All Rights Reserved.

G:

機会損失みたいなイメージですね。

鹿内:

そうです、機会損失がある状況ですね。

G:

データサイエンティストの需要に供給が追い付いていないという状態は続いているのですか?

鹿内:

はい。ますますデータ活用の範囲が広がっているので、データサイエンティストがいればいるだけ利益や売り上げにも貢献できる状況なのですが、人材がいないのです。もちろん数は増えてはいるんですが、それ以上に需要が高まっているのです。今までは、Web上で取得されていたデータを利用して、別の目的でも使うという方向でした。しかし、もっと積極的に、能動的に仕掛けていこうとなり、先ほどのブリヂストンさんのように、最初からデータを取りにいこう、使えるデータを取りにいこうという段階になっています。



G:

現状では企業はデータサイエンティストを外注している感じなのですか?

鹿内:

いえ、内製化する企業が増えています。一般に、先端的なことはベンダー企業と協働し、だんだんと自分たちだけでできるよう内製化していきます。データサイエンスにおいても、すでに内製化する企業は増えているのです。だから足りなくなっている。で、引く手あまたなので間に合わない。

G:

なるほど。企業側でも社内でデータサイエンティストを抱えたいという話になりますよね。

鹿内:

そうです。ユーザー企業は自社で持っているデータをきちんと分析できるようになります。これは受け売りなのですが、データサイエンティストは「掛け算の人材」だと思っています。今後はそういう人材の価値が高まっていくと思います。もちろん将来的に技術が洗練化されていくと、だんだん作業者になっていくものです。データサイエンティストもいつかは、ブルーワーカーになるか、人の職業ではなくなると思います。ただし、現時点では、「掛け算」で効いてくる人材です。生産性に掛け算で効く人材なので、少しでも良いデータサイエンティストを選ぶことが重要です。

G:

掛け算では1.2倍と1.3倍と0.1違うだけで、全然違いますよね。

鹿内:

全然違いますね。足し算の職種だと、2倍の生産性をもつような人材が1人いても、100人の部署を101人分にするだけです。1%しか変わらない。掛け算の人材だと、元々の売り上げが100億円もあると、1.2倍の人材でも120億円、20%も上がりますね。さらに、1.3倍なら130億円となります。しかも、1.2と1.3で0.1しか違わないのに、10億円も違うわけですよね。

G:

データサイエンティストの質でそんなに変わるものなのですか?

鹿内:

ものすごく変わると思います。古典的な統計のやり方で作った開発事例があったのですが、その開発を機械学習を利用して実施したら、開発期間1/3、開発者の人数も4割、費用は20分の1ほどで達成できましたで。良い人材、良いチームビルディングができると、費用も少なく、期間も短くなるという良い例です。いろいろな効率や価値を掛け算していくと100倍以上も違うことがあるのです。

G:

そういった100倍もの違いの大きさは、データサイエンティストの質で決まると。

鹿内:

質で決まります。先ほどのは上手くいった例ですけど、そうでなくても10倍くらいは変わると思います。



◆3:データサイエンスが使えないと給料が減る時代が来る!?

G:

データサイエンスの需要が高まっている、ましてや質の高いデータサイエンティストがまったく 足りていないという状況だと、企業も良い人が欲しい。そういう人材が欲しいという獲得競争が ある状況で、ぶっちゃけていうとデータサイエンティストは報酬面でめちゃくちゃ優遇されてい たりするのですか?

鹿内:

まだまだ、日本はこれからというのが現状ですね。というのは、業界によっては、そもそもデータサイエンティストがまだいない状況なので、彼らがどういう価値を生み出してくれるのかがわからない。

G:

そうですよね、そもそもデータを活用して利益が生み出せるという認識がないと、なかなか用意しようとはならないですよね。

鹿内:

企業としても、いない人材の人事制度はつくれませんからね。データサイエンティストの正しい

「価値付け」ができていないですね。私から見ると、上手く活用すればどの業界でも活躍するとは思います。あらゆる業種でデータサイエンティストは必要なのです。もちろん濃淡はあります。「こういうときには上手くいく、いかない」というのがある。それをちゃんと見極められると、高いお金が支払われるのですが、まだ日本ではそこに来ていないと思います。

G:

例えばアメリカだともっと上手く活用していますか?

鹿内:

そうですね。アメリカでも失敗はたくさんあると思いますが、雇用習慣の違いもあるからリスクをとりやすいと思います。日本だと特に大企業はデータサイエンティストを総合職で取ることが多いです。そのため専門職化がしにくくて、人事制度が一緒なんですよ。データサイエンティストだけを特別な給与体系にできないので、給与を上げることができないのです。

G:

なるほど。ただ、ITの技術やトレンド、ビジネススタイルはアメリカから流れてくるというか、ちょっと遅れて日本に導入されることが多いですよね。雇用体系についても日本は変わりつつあると思うので、今後は良いロールモデルとしてのアメリカを元にして、日本の人材活用がうまくなることは間違いない?

鹿内:

それは間違いないと思います。それでいくと、すでにヤフー株式会社さんなどのIT企業は専門職としてデータサイエンティストを雇っています。給与面での待遇の良さだけでなく社会人のまま博士課程にいけたり、学び続けるためのフォローもあります。他のIT企業では、国際会議に渡航費をバックアップして、就業時間の中で学べたり、スキルアップの機会を作ったりしているようです。このように企業によっては、すでにデータサイエンティストの待遇や学習支援を強化している企業もあります。

将来的には、データサイエンスは当たり前の技術になっているでしょう。今、オフィスワーカーがパソコンが使えないということはないですよね。そういった意味では、データサイエンスのスキルを持っているから給与や待遇が良くなるというよりも、スキルを持っていなければ、給与が下がると思った方が良いかもしれません。そのためにも学んでおくべきともいえるかもしれません。

G:

ちょっと怖い話ですね.....(笑)

鹿内:

もちろん、今は高い報酬も求めるべきところです。僕は10月で「期間の定めのない契約」の社員としては、パーソルキャリアを辞めました。11月から、パーソルキャリアの新規事業Data Shipと個別契約しています。データビジネスを作る事業開発者としてプロ契約を結んだわけです。パ

ーソルキャリアにあるミイダスともデータサイエンスプロジェクトを推進するために契約しました。プロのデータサイエンティストとして、次のフェーズに歩み出しています。単に契約期間に応じた報酬をもらうだけではなく、成果が出れば追加報酬が出るようなオプションを設定しています。新しい働き方を模索しているところです。

G:

今、データサイエンティストとして働いてる方は、どういう素養を持つ人が多いのですか?もともとはプログラマやSEだった人たちですか?

鹿内:

プログラマやSEに限らないと思います。それは多様ですね。マーケティングでのExcelやAccessでの集計やデータ分析からはじまって、データベースの整備の必要性が出てきて、システム基盤を地道に作っていった方の例もあります。特殊ですが、人事のデータ分析をしている女性の前職がネイリストだったという例もあります。彼女は文系だったようですが、よくよくヒアリングしてみると、記号操作など数学的な素養のある方でした。文系でも素養があるという意味では、心理学を研究していた方もそうです。心理学はたいてい文学部にあるんですが、統計科学の基礎はやりますよね。コテコテの文系に見られがちなんですけれど、研究でやっている人たちには統計の分析をできる人がいます。その人たちに、データサイエンティストとして活躍できると気づいてほしいです。

G:

データを取れないものって多分ないでしょうから、全ての分野で活用できるけれど、それを活用できると気づけないことがありそうですね。

鹿内:

心理学を学んだ人がどういう道に行くかというと、専門家としては、カウンセラーになったり、 教育の道に進む方も多いと思います。でも、少し視野を広げて、人を相手にする「人事」でピー プルアナリティクス(人事のデータ分析)をするキャリアも面白そうではありませんか?



G:

確かに人事というと、今までは感覚とか人間関係とかデータ化されてないものでフワッと決めが ちだった節がありますね。けれど、本当はデータを取って活用できればもっと良いマッチングが 生まれる、そういうチャンスがある領域なんでしょうね。

鹿内:

データがこれまで入ってないところこそチャンスなんですよ。先ほどの製造業もその一つです。 小売りなんかもPOSデータは使っているのだけれど、十分じゃないですよね。飲食店の予約シス テムなどもそう。トレタさんでは、来客の分析だけにとどまらず、座席配置の分析や自動化もや っている。飲食店で4人掛けのテーブルがあるけれど、「ここに2人入れていいのか、それとも4 人入れるべきなのか」とか。データ分析して、課題が見えて、改善できるというのは面白いと思 います。

G:

まったくデータサイエンスを活用していない企業も多い中、すでに活用してもっと積極的に育てていこうという企業もあるという状況なのですね。

鹿内:

今はちょうど産業構造の転換期だと思います。2000年代から始まってきたWebサービスが、10年から15年で一段落ついたところです。IT業界が、例えばGoogleがGoogle Glassを作ってみたり、バーチャルの世界ではなくてリアルな世界に踏み出そうとしています。Amazonも実店舗を作っています。さらにIoTの世界に移りつつあるところなので、私はデータサイエンスの活用をメーカー(製造業)にも頑張ってほしいです。製造業というモノが作れるところに情報学、データサイエンスが入ってくると相乗効果があるはずです。

G:

製造業はまだまだデータが活用されてない分野ということですか?

鹿内:

活用されていないですね。製造業のノウハウというのは簡単には真似できないので、その点では リードしているはずなのです。しかし、逆にいうとそこに情報学をまだ導入していない理由あっ たんですよね。IT業界は何も持たなかったから、ある意味、情報学とかデータサイエンスとかを どんどん取り入れてやってきたのですけれど。それでいうと製造業というのは、例えば自動車だ と機械系や化学系の人材がたくさんいるけれど、相対的に情報系は少なかったりします。

G:

これから自動運転カーの世界になると、さらにバランスが悪くなるというイメージですね。

鹿内:

いわゆるデータドリブンではないことも多いです。程度の問題ですが、モデルや知識を利用して数値シミュレーションしてはいるものの、論点の設計さえうまくやれば、もっとデータから成果を出せるのではないかと期待します。

iPhoneのSiriという音声認識がありますよね。10~15年前だと日本の音声認識の技術もトップ5には入っていたはずですが、Siriが出た当初は、まだそこまで実用的ではなかったと記憶しています。今だと結構、実用的に使えますよね?これはデータが貯まったから改善している部分が多くあると思います。サービスとしてとりあえずリリースをして、新しいサービスを面白がって使うイノベーターな人たちがいて、それによってデータが貯まる。データを蓄積するシステムを作れば技術革新が進む部分もあるので、とりあえず実施する。走りながら考えないと、上手くいかないだろうなという気がします。



G:loTであらゆるものが情報やデータを取れるようになれば、もっともっとデータが増えてきてもっともっと活用の範囲は広がりそうですね。



鹿内:

そうですね。そういう意味でいうとデータ分析よりも前の段階の「データ活用の設計」、つまり

ビジネスでどうお金になるかを考えていくことを、エンジニアやデータサイエンティストがやって欲しいなと思っています。そこに興味を持つ人材が増えるとうれしい。例えば、小売りの業界で、総合ディスカウントストアとして知られるドン・キホーテさんのデータサイエンスを担う株式会社リアリットさんは、新たなデータ取得に取り組んでいますね。そうじゃないと、今後の日本は危機的な状況になるかな、と。

G:

国際競争力で劣ると?

鹿内:

そうです。例えばGoogleとかFacebookなどいわゆる「GAFA」と呼ばれるところの創業者の多くは物理学でPh.Dを取っていたり、理系出身なんですね。日本では経営者で理系出身者ってそんなにいません。流石に製造業では、役員にも博士号取得者がいたりしますが、データ活用の専門家ではありません。科学や技術に関して高い専門性を持っている人たちがビジネスをちゃんと作る、お金に対しての責任を持ってやるというところの人材があらわれないとマズいだろうなと思っています。利益がでるほど、サービスが回り、データ(収集・分析)も回る状況なので、利益とデータは、もはや切り離せない状況です。そういうデータがあると技術革新も起こるし、相乗効果があらわれます。文系でもデータサイエンティストになれる一方で、大学院などで科学や技術のトレーニングを受けた理系は、データサイエンティストはもちろん、一定数は事業開発や経営にかかわってほしいですね。



◆4:「データサイエンティスト」になるための学習法とは?

G:

ここまでデータサイエンティストの仕事や現状について聞いてきましたが、では数が圧倒的に足

りないという「データサイエンティストにどうすればなれるのか?」という点についても教えて ください。

鹿内:

現在Data Shipでは、オフラインのイベントを定期的に開催し、様々な業種業態のデータサイエン ティストの講演、情報交換の機会を提供しています。オンライン学習も提供していて、Udemyは その1つとして推奨をしています。

Udemyは、米国法人Udemy, Inc.によって運営される世界で2400万人が利用するオンライン学習 プラットフォーム



鹿内:

Udemyは動画なので電車の中などどこでも見られるところが良いですね。Udemyには、とりあえ ず見る、というとっつきやすさがあります。他の勉強とも合わせて、復習に使うというのもアリ だろうし、動画の気軽さは圧倒的に大きいかな、と。本との比較でもそうです。本は持って歩く のも結構大変だけれど、Udemyだと多くの人がいつも持ち歩いてるスマートフォンでも見られま すから。

G:

「データサイエンティストになりたい」という人が具体的にUdemyで学ぶ場合、どういう風に学 ぶのが良いでしょうか。

鹿内:

プログラミングを学ぶのも第一歩です。

G:

データサイエンティストへの第一歩として、大前提としてプログラミングが必要だと。

鹿内:

学んだ方が良いと思います。実は、Data Shipは事業としてもインターンシップを受け入れているのですが、まずはExcelで作業してもらっています。2週間もすれば、いやでもその限界に気づきますよ。何度も同じ作業をすることになって面倒だと(苦笑)。CSVファイル、1つに収まる程度で、扱うデータは大きくなくてもそうです。面倒だと思ってくれれば成功です。

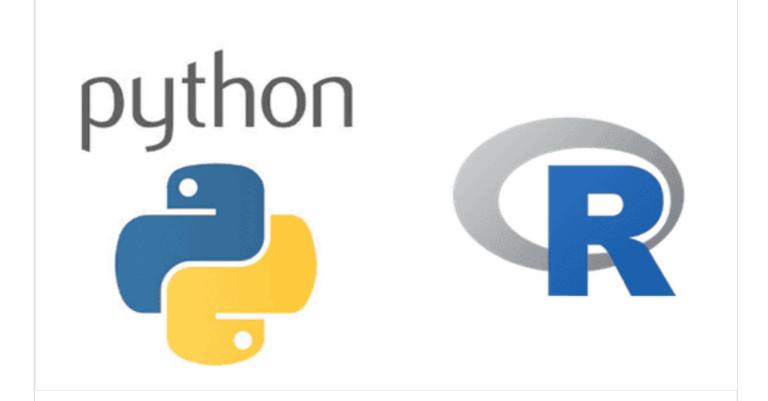
データサイエンスを学ぶ上で、言語はSQLとR、Pythonを軸にしています。SQLはデータベースからデータを取ってくるときに使う言語なので、これを使えるとデータ分析がすごくしやすくなる。反対にSQLができないと、データベースのエンジニアに頼まなきゃいけないので、時間がかかりますね。早く回せるようになるようにはSQLくらいは覚えておいた方が良さそうです。

G:

分析の言語としては、RとPythonをどう使い分ければよいのでしょうか?

鹿内:

エンジニアリングのシステムではPythonが使われだしていて、データ分析からシステム実装までが一気通貫できる言語として人気です。ただ、Rでやるのも手で、馴染みがある人が多いと思うので、Rも選んでいます。RもPythonも統計の「ライブラリ」がそろっているので、それはどちらでも良いかなと思います。



G:

分析の先のシステム開発まで見込めばPythonだったらスッと行くと。

鹿内:

そうですね。それとPythonのエンジニアは年収が高い傾向にあるようです。それも魅力の1つかと思います。

G:

プログラミングの知識を身につけていて、実際に仕事をしているような人がデータサイエンスを 独学で学びたい、やはり質の高いデータサイエンティストを目指したいという場合に、次に学ぶ べきことは?

鹿内:

データのことを学ぶために「可視化」をすること。表計算の数値の羅列にしないでおくこと、もっと厳密に言うと、データの分布を把握することです。次は、「そもそもグラフは、何の意思決定に使うんだっけ?」という、目的設計からデータサイエンスをきちんとやることですね。分析アルゴリズムを作って満足するという罠におちいらないことです。

その点で、Udemyの講座で**SIGNATE社(旧オプトワークス社)**が作ったものが良いなと思っています。そこでは具体的な事例が出てくるんですね。「何の目的のためにそれを作るんだっけ?」という部分がしっかりおさえられています。「どうやってデータを活用するか」というところを意識して学べるのがすごく良いと思います。それは多分、データサイエンスとエンジニアリングの大きな違いです。エンジニアリングは作ってみて、想定する動きかどうかはみるとわかりますよね。想定しない動きのテストは大変ですが。

けれど、データサイエンスは、アルゴリズムを作って予測してみると、とりあえず結果は出てきます。しかし、「間違った評価正解」をしてしまうことが多々あります。結果が本当に正しいかどうかを評価することも学ぶべきです。これは「バリデーション」と呼ばれる部分です。

例えば、「正解率で90%を達成しました!」というような発表がよくあるんですけど、「90%」を「精度の良い」と評価するのは、間違っていることも多い。例えば、人材業において「100人受けたら3人が採用される」というデータで「その人が受かるかどうかを予測しましょう」ということを目的にすると、実は97%で当たる方法がパッと思いつくんですね。それは、「採用されない」と答えるアルゴリズムです。この場合、ほとんどの人は採用されないので、「採用されない」と予想すれば、100人中97人は当たるんですよ。

G:

でも意味ないですよね?

鹿内:

意味ないですよね。採用されるか、されないかなので、五分五分、50%にみえますけど、1つ知識を入れると97%は絶対当たるんです。けれど、これでは予測してないのと一緒でしょう。この97%から98%、99%にするところが本当は勝負所なのです。単に「97%で当たりました」と言われても、それでは何の意味もないのですよね。それに気づかずにプレスリリースとかで発表しているケースが結構、多いです。



G:

数値としては高いのだけれど、実体を反映していないと。

鹿内:

何の意味もないですね。それを考慮して評価をしないとダメです。エンジニアの人はその観点を 見落としがちなので、バリデーションを意識すると良いデータサイエンティストになれると思い ます。

G:

そういうものって学習できるものですか?

鹿内:

もちろんです。僕も最初は20代の最初は知りませんでした。ちゃんとそのことに触れているデータサイエンスのコースもあります。

G:

データサイエンティストになるための背景事情は多様だということですが、現状では「一から独 学で始めて、質の高いデータサイエンティストになる」ということは可能でしょうか?

鹿内:

目指すレベルにもよりますが、独学だけでは難しいでしょう。会社のOJT(On the Job Training)だけでも難しい。会社では「経験10年の人」というような呼ばれ方をしますよね。これは職務経験でスキルを測っているのです。けれど、企業でデータサイエンスの経験10年の人は、なかなかいませんよね。

大学や研究機関では、若い大学院生たちが世界の最前線でデータがかかわる研究プロジェクトを やってるわけですよ。データサイエンスの世界では若くてもスキルフルな人がいっぱいいます。 その経験さえあれば、職務経験や仕事のビジネスの経験がなくてもすごく使えるので、私はそう いう人も推薦します。

G:

要するに経験がなくても即戦力になれるということですか?

鹿内:

はい。経験は、「職務」経験だけではないということです。逆に20年前に大学を出た人の場合、企業に入ってからも論文を読む習慣がなければ、最先端のことをキャッチアップできていないはずなので、そういう意味では不利になっているはずです。それこそUdemyでもう1回、基礎から復習し直すことなどが必要だと思いますね。

G:

常に学習し、自分で追いつく努力をしていないとあっという間に取り残されるんですね。

鹿内:

データサイエンスに関わらず、おそらく今の時代は技術発展が圧倒的に早い。昔だと大学でやってたことが20年ぐらい経つとビジネスサイドの最先端として降りてきていたのですが、私の感覚だと、データサイエンスは10年くらいで降りてきたな、という感じがしています。それはやっぱりビジネスサイドでデータがたくさん貯まっていたというのもあります。大学だけじゃなくて、一部の研究ではビジネス側の最先端にもなっているんですよね。データサイエンスは、アカデミックとビジネスがくっついてる状況だと思います。

G:

じゃあもうアカデミックの技術とかをいかに早くビジネスに持ってくるのかが大事ですね。

鹿内:

大事です。技術を扱う人材という意味でも、大学院生は重要です。大学院生のように「2~5年

間、データ分析のかかわる研究をやる」ということは企業でなかなかできない。

もし、いま、データサイエンスの仕事に関わっていなくても、2年前に大学院で研究していたなら、Udemyの機械学習の講座を見てもらいたい。データサイエンスの素養をはぐくむ教育をうけているはずなので、「あれ、自分にもできるんじゃないか?」と感じられるはずです。最前線を離れていたというのであれば、まずは復習がてらにUdemyの動画講座に取り組んでやり直す。そうすると、別のキャリアを歩み始められるかもしれません。

◆5:プロがおススメするUdemyデータサイエンス講座5選

G:

最後に、鹿内さんには「これからデータサイエンティストになろうとする人」のために、オスス メ講座を事前に5つ挙げていただきました。この5つの講座のメリットをそれぞれ簡単に教えてく ださい。

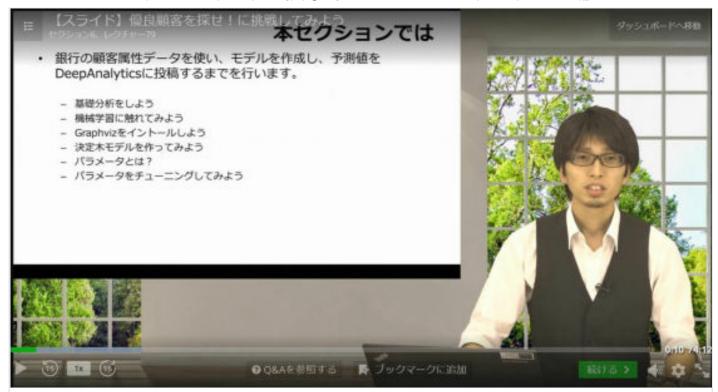
◆1: 【ゼロから始めるデータ分析】ビジネスケースで学ぶPythonデータサイエンス入門



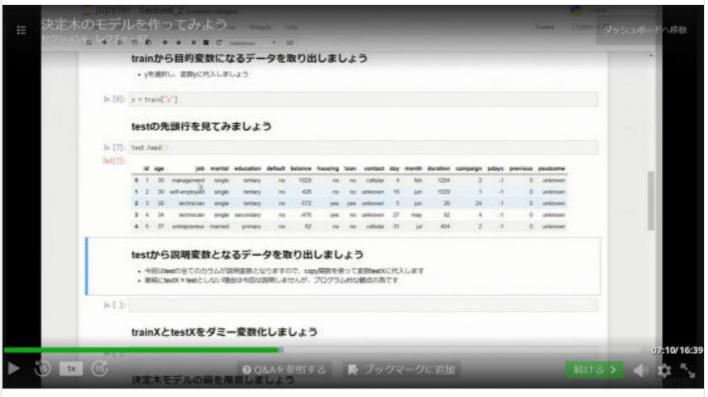
鹿内:

「【ゼロから始めるデータ分析】ビジネスケースで学ぶPythonデータサイエンス入門」は個人的に最もオススメできます。プログラムの環境を作るところから入ってる点も良いですね。プログラムの開発環境を作るところっていろんなインストールが必要なんですけど、意外とつまづくし、面倒なんですよ。プログラムをやったことのない人が一番つまづくとこかもしれません。

また、「弁当の売上予測」「銀行の顧客ターゲティング」という2種類のビジネスケースで構成されているので、プログラマ、エンジニアではない人たちでも興味を持てると思います。



身近なケースで興味を持てればデータ分析を進められるような内容になっているし、バリデーションまで含まれているので、データを取得した後の工程は全部入っています。企業でのデータサイエンス業務を一通り体験できます。



あとは、初学者が取り組むべき必要最低限のことに絞られている点も良いです。情報過多だと、混乱しますからね。回帰分析とかクラス分類の一番使うアルゴリズムが紹介されています。多くの企業でいうと95%くらいは、それで済むのじゃないかな。今はディープラーニングがホットだということで、エンジニアはみんなそこに食いつくのですが、ディープラーニングを適用できるデータをもつ企業は少ないと思います。本当にIT企業でデータが大きいという条件であれば良いのですが……。線形回帰などの、まさしくこれを最初に使う、というのが解説されているのが良

いです。

◆2:はじめてのSQL・データ分析入門 - データベースのデータをビジネスパーソンが現場で活用するためのSQL初心者向コース

はじめてのSQL・データ分析入門 - データベースのデータをビジネスパーソンが現場で活用するためのSQL初心者向コース

エクセル作業からのステップアップ!データベースのデータを取得し、効果的 にレボーティングしたいビジネスパーソンにおすすめ。データベースを操作す る言語SQLを短期間で身につけビジネスの現場で使うための実践型カリキュラム。MySQL使用。

ベストセラー | 青 青 青 青 青 4.3 (224件の)評価) 1784人の受罪生が登録

作成者 中村補太 (Yuta Nakamura) - 最終更新 11/2018 - ● 日本語



¥ 1.200 +49,800 94%OFF

カートに入れる

今すぐ購入

30日間返金保証

その他:

- 7時間のオンデマンドビデオ
- 11件の記事
- 会 学習期間の展開なし
- □ モバイルとPCからアクセス
- 拿 极了证明售

クーボンをお持ちですか?

学習内容

- ✓ SQLを使用してデータベースからデータを 取得する
- SQLを使用してデータ分析を実行する
- データベース構造の変更ができる。
- ✓ MySQL公式ツール MySQL Workbenchが使用できる
- √ データベースのデータを更新できる
- ✓ MySQLをmacOSまたはWindowsにインストールできる
- ✓ MySQLを操作できる

◆3: 超実践!「Rで学ぶビジネスデータ分析」講座

超実践!「Rで学ぶビジネスデータ分析」講座

「R言語」を手で動かしながら、実務で使われている「データ分析」(データサイエンス)の「4つの手法」を習得します。ぜひ! 実際のビジネスの現場で活用されている、実践的な「ビジネスデータ分析」(ビジネスアナリティクス)を手に入れてください。

★★★★ 38 (91件の評価) 1151人の受講生が登録 作成者 Ichiro Takahashi 最終更新 2/2018 ● 日本語

このコースをプレビュー

¥ 1,200 ¥19,800 94%OFF

学習内容

- 実務のビジネスデータ分析(ビジネスアナ リティクス)の考え方や流れ、手法などを 理解できるようになる
- ピッグデータを使った、営業生産性や販売 効率を高め方が、どのようなものか分かる ようになる
- ✓ ビジネスデータ分析で使う定量分析を、R
 で実践できるようになる。
- √ データを含めた情報活用のPDSサイクル (or PDCAサイクル) とOODAループの造 いが理解できるようになる

カートに入れる

今すぐ購入

30日間遊金保証

その他:

- ② 10 時間のオンデマンドビデオ
- □ 3何の発薬
- ② 23個のダウンロード可能なリソース

鹿内:

「はじめてのSQL・データ分析入門」と「超実践!『Rで学ぶビジネスデータ分析』講座」も、ベーシックかつ、初学者に優しい「過不足ない構成」のカリキュラムで構成されています。初学者は、多すぎても情報に溺れてしまうし、足りなければ上手く活用できないので過不足ない構成

というのは重要です。

以上の基礎となる講座に対して、次の2つはアドバンスドな内容が学べるので選びました。

◆4: 【TensorFlow・Keras・Python3で学ぶ】時系列データ処理入門(RNN/LSTM, Word2Ve c)

【TensorFlow・Keras・Python3で学ぶ】時 系列データ処理入門 (RNN/LSTM, Word2Vec)

TensorFlow, KerasとPython3を使って、自然言語処理や時系列データ処理を学 びましょう。日本語+動画で学べる唯一の講座(2017年8月現在)です。 RNN/LSTMは、機械翻訳、自動字幕表示、株価予測などに使用されています。

📷 ★★★★ 4.1 (381件の評価) 3833人の受護生が登録

作成者 井上 博樹 (Hiroki Inoue) - 最級更新 3/2018 - 🗩 日本語

学習内容

- ✓ Word2Vect

 Sequence2Sequenceなど TensorFlowの時系列データ処理のチュート リアルを理解できるようになります。
- RNNの動作原理について理論を学習し、コ 一ドを思いて理解を深めます。
- ✓ Kerasを使用したLSTMによる文章合成プロ グラムを作成できます。
- ✓ TensorFlowを用いたRNNによる機械学習 ができます。
- ✓ LSTMの仕組みについて理解し、コードを 実装して理解を深めます。



カートに入れる

今すぐ解入

30日間提供金保証

その他:

- ☑ 6時間のオンデマンドビデオ
- □ 9年の記事
- 部間のダウンロード可能なリソース
- 会学管期間の制限なし
- □ モバイルとPCからアクセス
- 康 修了证明男

◆5:みんなのディープラーニング講座ゼロからChainerとPythonで学ぶ深層学習の基礎

みんなのディープラーニング講座 ゼロから ChainerとPythonで学ぶ深層学習の基礎

初心者向けディーブラーニングのコースです。ニューラルネットワーク用ライ ブラリChainerとプログラミング言語Pythonを使って、深層学習の基礎を習得 しましょう。人口知能 (AI) や機械学習の分野で、特に注目を集めている技術 です。

ベストセラー 4.1 (287件の評価) 2755人の受講生が登録

作成者 我妻 幸長 (Yukinaga Azuma) 最終更新 10/2018 ● 日本語

このコースをブレビュ

¥ 1,200 +15,000 92%OFF

カートに入れる

今すぐ購入

30日間海金保証

その他:

- ② 5時間のオンデマンドビデオ
- ☐ 6FF/DBD#
- ② 2億のダウンロード可能なリゾース
- 学習期間の無事なし
- モバイルとPCからアクセス
- **商 位了证明者**

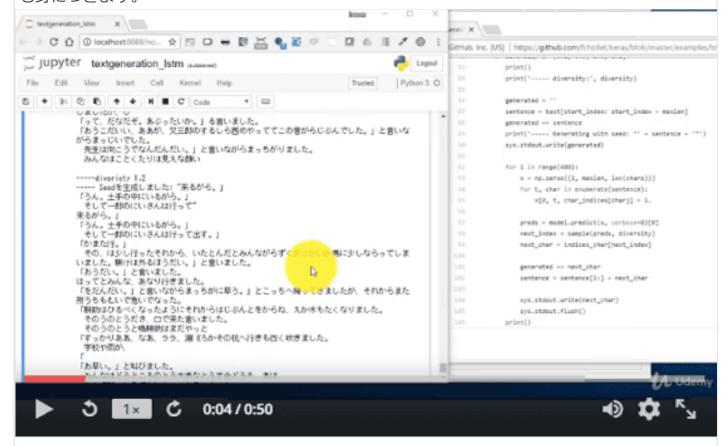
学習内容

- ディープラーニングの仕組みが身につきま
- ✓ ライブラリChainerの扱い方が身につきま
- ✓ 畳み込みニューラルネットワークの基礎が 身につきます。
- ティープラーエングが持つ可能性に気づ き、人工知能に対する深い消費ができるよ うになります。
- ✓ Pythonの基礎的なプログラミング技術が食 につきます。
- ビジネス上必要なディーブラーニングの差 礎知識が身につきます。
- ✓ さらに高度な応用を行うための下地が身に つきます。

鹿内:

「【TensorFlow・Keras・Python3で学ぶ】時系列データ処理入門(RNN/LSTM, Word2Vec)」では、具体的には、高度なアルゴリズムを作成するのに便利なライブラリーの使い方を学ぶために、Googleが開発した「TensorFlow」を学習できます。また、ニューラルネットワークのモデルを作るときに便利な「Keras」の使い方も学べます。さらに2つを利用しながら、word2vec、RNN、LSTMというアルゴリズムを、機械翻訳、感情分析、株価予想といった具体例で学べます。ライブラリーのインストールなど開発環境構築がWindows・macOSの両方で解説されていてとても親切です。

「みんなのディープラーニング講座」は、日本が誇るAIスタートアップであるPFNが開発したライブラリーChainerを利用しています。Pythonについても、「Hello world」から解説があります。エンジニアであれば、いきなりこの講座から始めても良いかもしれません。バックプロパゲーション、活性化/損失関数、勾配降下法など、ニューラルネットワークに関わる数理的な概念も身につきます。



鹿内 .

データサイエンスが、エンジニアリングと一番異なるのが、予測精度の評価なので、Udemyでの学習に加えて、仮設検定やクロスバリデーション、情報量基準など、予測精度の評価をするための概念や方法についても学んでいくことをオススメします。

G:

なるほど。本日はいろいろと貴重な話をありがとうございました。

◆Udemy秋の大セール(Black Friday & Cyber Monday Sale)

インタビューで出てきたオンライン学習プラットフォーム「Udemy」で、ブラックフライデー&

サイバーマンデーセールが実施されています。2018年11月28日17時まで最大95%オフで受講ができます。新しい知識や技術を身につけたい人、スキルアップを狙う人には、絶好のチャンス到来です。

なお、オススメのUdemy厳選10コースの一挙レビューは以下の記事で確認できます。

【最新講座多数】AWS・AI・ブロックチェーンなどの技術が動画で学べるUdemy秋の大セール 【11/28 17時まで最大95%OFF】 - GIGAZINE



1 プログラミングをあな*1* に

プログラミングはもはや必須のスキに。これからの世界を生き抜くた&今すぐ身につけましょう

テックキャンプ

2 役員経験がある方必見

顧問となり月数回の会社訪問で、現時を超える報酬を。こちらから稼働件をご確認下さい。

レイスMS株式会社



関連コンテンツ



機械学習やビッグデータを 扱うデータサイエンティス トの年収や使用言語などを 赤裸々にするデータ



くの学生が仮想通貨とフ ロックチェーン技術に興味 を持ち始めている



SEMを得意とするネット広 告代理店「アイレップ」に 行ってきました



「生涯現役開発者」を貫く 40代~60代のソフトウェア 開発者たち



Google会長が「人工知能は 世界に横たわる厄介な問題 の解決にこそ役立てられ る」と語る



人工知能による自動化が進 むゴールドマン・サック ス、人間のトレーダーは600 エンジニアやプロダクトマ 人から2人へ



これがIT業界の給与事情、 世界各都市のソフトウェア ネージャーなどの平均給与 を比較するとこうなる

<< 次の記事



検索エンジンの「Google」にコメント機能が追加さ れる

前の記事 >>



誰かをハグすることでストレスが軽減されるという 研究結果、ただし合意に限る



2018年11月21日 09時00分00秒 in 取材、インタビュー、ソフトウェア、広告, Posted by logy to

世界で4番目に幸せな国「バヌアツ共和国」には人喰いの村がある? ウワサの真偽を確かめてきた【インターリンクドメイン島巡り第9回「.vu」】



広告

「GIGAZINEマンガ大賞」2019年2月度募集開始&トップ画メイキングはこんな感じ



マンガ



最新ニュース40件

- パーキンソン病のパンデミックが将来起こる可能性があると研究者が警告
- 実写とCGIを完全融合させた「アリータ:バトル・エンジェル」のメイキング映像公開
- 「YouTubeが子どもの性的搾取を促進している」と批判の的に、ディズニーやEpic GamesがYouTubeへの広告を引き上げる事態に
- EUの改正著作権法指令案第13条はYouTubeのコンテンツやクリエイターにどのような影響を与えるのか?
- 低カロリーなミスドのドーナツ「もっちりフルーツスティック」は健康とおいしさを両立しているのか?を 食べて確かめてみた
- 美少女イラスト風に自動生成された「俺の嫁」画像をズラッと大量に並べて見せてくれる「These Waifus Do Not Exist」
- レオナルド・ダ・ヴィンチが「理想郷」としてデザインするはずだった幻の首都がある
- 「キングダム ハーツⅢ」に熱中するあまり「幸運のマーク」を現実世界でも探し求める人が続出
- 2019年2月22日のヘッドラインニュース
- 16言語を独学で習得した同時通訳者ロンブ・カトーに学ぶ語学学習の極意とは?
- 30年以上目撃されていなかった世界最大のハチが探検隊によって再発見される
- この世に存在しない「架空の猫」の写真を自動生成しまくれる「This cat does not exist」
- トランプ大統領がこの世に存在しない「6G」を要求していると話題に

- テスラのModel 3が権威ある消費者信頼性調査の「オススメ」から漏れる
- 「ゾンビ鹿」の肉を食べると人はいったいどうなるのか?
- Facebookはナチスや白人至上主義に興味を持つユーザーを特定してターゲット広告の表示に利用している
- 大量の蒸しあさりとトビウオのダシで磯の香り凝縮されたリンガーハット「あさりたっぷり春ちゃんぽん」 を食べてきた
- AppleがiPhoneと連携させて使う独自のクレジットカードを準備中か、Apple製品の購入時は高キャッシュバックが受けられる可能性も
- 「任天堂の顔」として世界的に知られるレジナルド・フィサメィ社長が2019年4月での引退を発表
- Appleも折りたたみiPhoneを開発中、どんなデザインになるのか?
- Intelは「Appleは2020年にMacをARMプロセッサに移行する」と予想している
- 消費者の商品選択を誘導する「おとり効果」とはどのような販売戦略なのか?
- 学生の課題の採点をGitHubとTravis CIで自動化することで成績向上や採点にかかる時間短縮などの利点が得られる
- ストレージの使用状況をグラフ化してどのファイル・フォルダが容量を無駄に使っているか一発で発見できる「Folder Size」
- 人の性的指向を変える「転向療法」は自殺や自傷を引き起こすと調査で示される
- 赤い麺と黒いスープでゴジラの世界観を再現したという「ゴジララーメン」を実際に食べてみた
- OpenAIが「危険すぎる」として文章作成AIの論文公開を延期したのは間違いだったと研究者が指摘
- 盗んだiPhoneのアクティベーションロックはどうやって解除できるのか?
- 禁欲的な教育は思春期の出産を減らさないどころか増加させる
- 「銃夢」の作者・木城ゆきと氏にインタビュー、「アリータ:バトル・エンジェル」映画化の経緯から「銃夢」の奇想天外なストーリーの秘訣まで聞いてきました
- 2019年2月21日のヘッドラインニュース
- 「イップ・マン外伝 マスターZ」ミシェル・ヨーの美しく力強い戦いをまとめたメイキング特報第2弾公開
- 「アリータ:バトル・エンジェル」プロデューサーのジョン・ランドー氏にインタビュー、制作経緯や技術の進歩など色々と聞いてみた
- 「ネズミの鳴き声の意味」を機械学習で解読するシステム「DeepSqueak」が開発中、マウス実験で活用される見込み
- 「ロード・オブ・ザ・リング」ドラマ版の脚本は厳戒に警備された密室で執筆されている
- マイケル・ジャクソンに性的虐待を受けたという告白を追うドキュメンタリー「Leaving Neverland」
- 濃厚アボカドを軸にした「アボカド生ハムモッツァレラバーガー」&「アボカドチリチーズバーガー」をウェンディーズ・ファーストキッチンで食べてきた
- シマウマがしま模様なのはなぜか?という理由がついに解明されつつある
- Samsungが折りたためるスマホ「Galaxy Fold」や「Galaxy S10」シリーズを発表
- 老舗の圧縮解凍ソフト「WinRAR」に19年以上前から存在する脆弱性が発覚





Surface Go

新しい Surface Go は、Windows 10 (S モード)¹ のパワーを活用し たノート PC ...

¥98,064

今すぐ購入 >

過去の記事

<		>				
日	月	火	水	木	金	±
27	28	29	30	31	1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	1	2

年と月を選択

▼



Surface Go

新しい Surface Go は、Windows 10 (S モード) 1 ...

¥98,064

今すぐ購入 >



カテゴリ

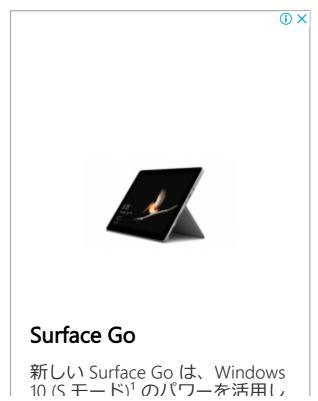
ヘッドライン	レビュー
取材	インタビュー
試食	モバイル
ソフトウェア	ネットサービス
ウェブアプリ	ハードウェア
乗り物	サイエンス
生き物	動画
生き物	動画
映画	動画 マンガ
映画	マンガ
映画アニメ	マンガ ゲーム

セキュリティ	お知らせ
ピックアップ	コラム
メンバー限定	広告

サイト内検索







たノートPĆ

¥98,064

今すぐ購入 >

人気記事ランキング

直近24時間に読まれた記事のランキング

- 1. 「ゾンビ鹿」の肉を食べると人はいったいどうなるのか?
- 2.30年以上目撃されていなかった世界最大のハチが探検隊によって再発見される
- 3. 「ブドウを電子レンジでチンするとプラズマが発生する」という現象の原因がついに明らかに
- 4. 2019年2月22日のヘッドラインニュース
- 5. トランプ大統領がこの世に存在しない「6G」を要求していると話題に
- 6. 2019年2月21日のヘッドラインニュース
- 7. Intelは「Appleは2020年にMacをARMプロセッサに移行する」と予想している
- 8. 16言語を独学で習得した同時通訳者ロンブ・カトーに学ぶ語学学習の極意とは?
- 9. Appleも折りたたみiPhoneを開発中、どんなデザインになるのか?
- 10. 「任天堂の顔」として世界的に知られるレジナルド・フィサメィ社長が2019年4月での引退を発表

人気動画

2019-02-16から2019-02-22の再生回数ランキング







ハイランド「ド・ドドンパ」の加 速力3.75Gをプレス向け試乗会で 体験してきた



動作ムービー





「DAEMON X MACHINA(デモン ポケットモンスター Let's Go!の エクスマキナ)」で高速メカ戦闘を 「GOパーク」で25匹のピカチュ 楽しむとこんな感じ

ウを集めてみた



誤字脱字の指摘	ネタのタレコミ	熟練のハンターがイノシシを解体
その他のお問い合わせ	広告掲載について	して内臓を取り出しているところ
GIGAZINEについて	採用情報	