

# AI Prediction & Ecosystem White Paper <sup>CN</sup>

AIPE COIN:  
TABLE OF CONTENTS

---

摘要	01
服务项目	02
3PIKS	02
AI比特蜂 (AIBitBip)	02
AI自动型交易模式 (AI Auto-Trading)	02
AI预测平台 (AI Prediction Platform)	03
向交易所提供解决方案	03
AI预测生态体系	04
AI技术	05
概述	05
主要技术介绍	06
通证经济体系	14
加密货币介绍	14
分配内容	14
款项运营	15
路线图	15
通知及注意事项	16

# 摘要

“人工智能”（Artificial Intelligence；简称AI）曾经以一场堪称围棋人机大战而轰动了全世界。在本次对战当中，阿尔法围棋（AlphaGo）与李世石九段之间的较量也随之引来了各界的广泛关注。在这之后，人工智能技术也开始被各个领域备受瞩目。从金融相关的服务及保险领域，到生命科学与教育研究领域也在积极地使用该技术。不仅如此，人工智能技术还被使用在政府、国防、制造业、能源、运输、医疗等领域中，还逐渐扩展到了销售、通信、传媒、娱乐、体育等生活领域。

实际上，随着人们开始对智能型商务模式的实际需求不断变大，对于传达到消费者群体所接受的经验值也不断被推陈出新。为了满足这一变化，谷歌（Google）、微软（Microsoft）、IBM、亚马逊AWS、英特尔（Intel）、FICO（Fair Isaac Corporation）、Salesforce、百度、SAS等全球知名企业也开始致力于把人工智能技术与商务应用两者进行相结合。

根据全球知名市场研究机构 Research and Markets 的发布的报告显示，从 2018 年 7 月到 2023 年截止，有关人工智能市场规模预期每年平均增长 48% 以上。机构能够获得该相关数据，主要原因在于有关加强智能型商务程序的必要性，物联网（IoT）、连接设备（Connected Device）的增加等情况而导致的大规模数据生成，以及后附加的运营费用削减的必然趋势等起到了重要作用。不仅如此，对于提高基础设施及数字城市相关项目的水平，全世界的各个政府，企业，以及地方自治团体等相关组织与机构也正在不断努力，并促进了人工智能以及相关服务市场的快速发展。该机构还将预测相关市场的发展将会持续一段时间。

人工智能技术是通过可持续发展，能够生成高附加值的一项产业。同时，它也是能够创造出实际收益的一个巨大市场。但是，对于该市场向外部所呈现出的发展趋势以及发展潜力相比，对于该技术能够解决市场上所发生的问题处理能力，现还处于初步的阶段。

对此，我们 AIPE 以人工智能预测（AI Prediction）为一个核心点来搭建一项人工智能模型。我们可以通过相关数据来进行测试，并将此技术实际应用到商业化当中的人工智能研究人员提供相关服务。不仅如此，我们也将根据人工智能模型所呈现出的数据，对加密货币进行投资的投资者之间提供可联系的一项平台。我们将利用人工智能开发者以及该团体所积攒的相关人力资来研发出一项以区块链技术为主的平台。该平台能够向需要获得人工智能数据的企业提供联系与服务。最后，我们将通过该项目，希望能够建立一个研究者与投资者以及企业，三者之间的紧密且互利的循环模式关系。

对于该平台内所产出的并共享的所有信息，我们将通过区块链技术来进行安全管理。人工智能研究人员对于所开发出的相关内容持有版权与管理权限。对于使用人工智能模型的投资者而言，个人信息以及通过该模型进行参考并进行投资等也将获得保护。不仅如此，所有的参与者与企业将通过该平台的共享而起到积极的作用，均可获得相应数额的奖励。

所有的参与者以及企业将由 AI 预测平台所提供的应用程序编程接口（API），在此平台内构建多样化的点播服务（On Demand Service）。我们的 AIPX 团队是通过人工智能预测平台（AI Prediction Platform），由人工智能研究人员所开发的预测模型（Prediction Model）与实际用户们所拥有的所有权与管理权限等进行有效再分配，给予相应的奖励。我们希望以此来创造价值，并形成一条良性循环的闭合。并且，以此作为基础而实现人工智能产业的创新，且对相关领域给予贡献。

# 服务项目

## 3PIKS

---

### 结合人工智能与区块链两项技术的全新的奖励平台

我们为了能够向市场提供以“人工智能”与“人为因素”(Human Factor)两者相结合的加密货币预测服务，专门开发了通过人工智能来预测加密货币市场的一项3PIKS服务平台。

对于加密货币市场的参与者而言，3PIKS服务中所提供的人工智能模型的价格预测结果以及3PIKS用户们所预测的结果均可作为其参考的主要数据指标。用户们将根据由我公司所开发出的人工智能模型一起预测加密货币的实际价格，以此作为奖励来获得AI Token。对于AI Token，未来可以与AIPX Coin进行相互兑换。不仅如此，AI Token还对即将要建立的战略伙伴公司所推出的通证及加密货币可以进行兑换。我们所构建的生态体系也是通过可在多处不同地点使用相关加密货币等来不断得到完善。

## AI比特蜂 (AIBitBip)

---

### 以人工智能为基础的加密货币价格变动预测信号服务

该服务的设计是以人工智能模型专门对实际的交易提供相应的帮助为主要目标。对于交易时可能会出现需要关注的情况（价格的上升与下降）下，该人工智能将提前收集相关数据，以发送信号的方式进行相关的通知服务。为了能够使用户能够更加方便地使用到该技术，我们将以手机软件(APP)的形式来开发样机，并计划在2019年第二季度正式发布。在初期，我们将优先使用由我公司自行开发的数十个人工智能模型，并进行该服务的提供。之后，我们打算通过AI预测平台中所开发出的诸多人工智能模型当中进行筛选以后，在此基础上对相应的模型进行内容追加与技术的扩张，并正式公之于世。

AI比特蜂 (AIBitBip) 所拥有的价值不仅限于提供加密货币交易的相关信息。不同的群体通过该技术所提供的信号而进行相关交易，将会形成一个巨大的价格主流群体。如果在同一时间段内，投入大量资本，将会形成一个总体价格的重要趋势。这也在一定程度上可以大幅度地降低相关投资的风险。因此，用户可以对该数据及价格趋势作为投资的相关参考资料，我们也很期待用户的实际投资风险也随之降低。手机软件的特点是方便，灵活性强，易于推荐给其他更多的用户。这有利于在软件市场内进行各种形式的广告发布与推广。

## AI自动型交易模式

---

### 以人工智能预测信号 (AI Prediction Signal) 为基础的自动交易模型 (Auto-Trading Bot)

即使，以AI比特蜂 (AIBitBip) 提前收到了预测信号，不是每一个人都可以在交易当中受益。在24小时内不停

地进行大规模同步交易的加密货币市场当中，凭借以提前收到的信号为来实际地进行交易是一件相当困难的事情。为了解决这一问题，我们将通过各个交易所的API来提供AI自动交易模型的相关服务。为了实现该项目的成功落地，我们已通过人工智能预测信号（AI Prediction Signal）来进行了实际的交易。虽然，在这期间我们也曾出现过多次错误与失败。但是，与此同时，我们也发现人工智能不断对相关过程进行学习后，对其预测的准确率及回报率开始回升并显示出上升趋势。

我们的服务是能够向所有的用户提供可以在正确的时间点上及时进行交易，且从而获得相应收入。对于使用该服务的用户须持有一定数额以上的AIPE，对于自动型交易模式中所获取的一部分收益也将与研究人员一同共享，以此来构建并维持整个生态体系。

而且，我公司所推出的自动型交易模式（Auto Trading）不仅限于只获取收益，而是能够通过该平台与交易所建立紧密的合作关系。有关进行该自动型交易模式所提供到的服务内容，我们需要不断的重复以进行出价登录以及成交的整个流程。通过该流程，我们可以确保交易所的更多交易量与加密货币的流通性。而且，对于在形成相应的市场，该服务平台将会起到重要的作用。我们期待通过B2B的商务模式，能够形成以用户，交易所，预测平台这三者之间的相互发展，并形成一个完整的生态体系。

## AI预测平台

---

### 与可确保已验证完毕的数据进行确认的人工智能预测平台 (AI Prediction Platform)

现阶段，对于人工智能研究中最大的技术瓶颈无非是对于大规模数据的提炼与计算能力的提升。以个人来对整体环境的设定与安排几乎无法实现，对于需要获取人工智能模型的相关数据的企业也仍然面临着严峻的考验。

AIPE是为了解决这一问题，并提高改善人工智能研发环境而推出的AI预测平台。在该平台内，人工智能研究人员可在开源图书管理系统环境下轻松便捷地使用该服务，同时企业也可以确保相应的使用人数（Traffic）以及模型数据的获取。在这种环境下，如果诸多人工智能研究人员可通过平台所提供的数据工具来搭建多样化的人工智能模型，这将对提供云端的相关企业在进入到市场内部时所碰到的壁垒也将可大幅度降低。如使用该模式及平台，对于AI模型的管理以及数据的共享将会变得更加容易。不仅如此，对于减少使用服务器的费用，增加多样化的交易模型或算法，以及增加对于用户的方便使用等优势可将带动对于新型资产运营模式的主流趋势。

综合以上内容，AIPX所提供的基础架构平台（Platform Infra）不单单只提供对AI研究人员与投资者，以及企业的方便服务。而是，对于能够带来人工智能产业当中的整体价值链（value chain）以及赋予全新理念的新思潮有着重大意义。

## 交易所解决方案

---

现有很多企业不断地向相关加密货币交易所研发以及提供其相关服务等，作为目标进军到该市场当中。但是，碍于时间与实际投入到的费用，加上专业技术的不足，加上营销战略的缺乏等，从而出现了导致诸多方面问

题及阻碍。

AIPE为了解决这一问题，我们已计划提供由我公司所开发的加密货币交易所与解决方案。以往的传统概念的中心化交易所出现的问题，我们也可以提供相应的服务。不仅如此，我们正在开发以加密货币为中心的，可进行保证金以及杠杆交易的保证金交易所(Margin Exchange)。我公司所开发的该交易所与BitMEX一样，是可以提供期货保证金的交易，在此还计划增加对于保证金的期货交易。

现阶段，可以提供有关加密货币交易所相应解决方案的企业数不胜数，但却仅限于以中心化交易所模式来提供相应的解决方案。AIPE是可以提供以保证金交易为主的交易所，还可提供作为衍生产品的期货交易，加上以提供期权交易为主，同时也可以追加保证金交易的保证金期货交易。不仅如此，我们还计划提供可支持保证金期权交易的加密货币投资解决方案。

## AI预测生态系统

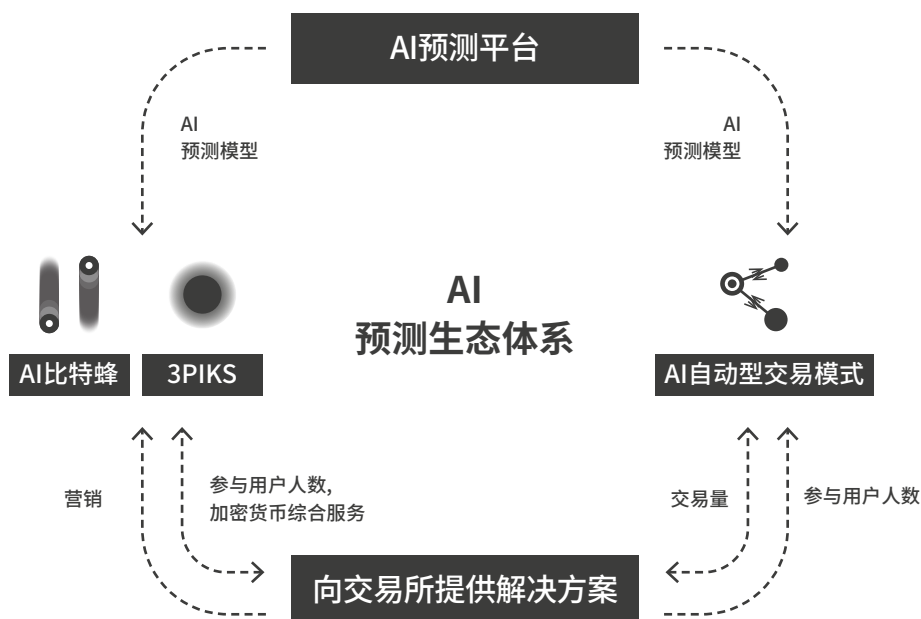
**AIPE所展望的是以人工智能为基础的，由各项服务所构成的生态体系。**

AIPE所追求的目标是能够使各项服务的价值体现形成一个有机的加密货币生态体系。

作为奖励平台的3PIKS与提前发出加密货币价格变动信号服务的AI比特蜂，通过使用这两项服务，用户就可以对加密货币或通证进行空投，也可以作为以交易所内的特定加密货币的营销渠道来使用。

不仅如此，现3PIKS内正在进行作为奖励而获得的AI Token(暂定)，并且能与签约合作方的通证或加密货币可进行兑换(Swap)的活动。对已进行兑换(Swap)完毕的加密货币，用户可以通过由我公司提供营销服务的相关交易所进行提款，并确保交易所的参与用户人数与流动性。

AIPE的经济系统不仅限于获得优秀的人工智能模型，也不仅限于只提供在开发交易所时需获得的解决方案。各项服务都互相保持一定的关系并构建有机的价值体系。同时，它也将形成加密货币价格预测，货币的流动性，宣传，参与人数以及可以保持市场稳定性等诸多因素结合在一起的一个整体的循环机制。



# AI技术

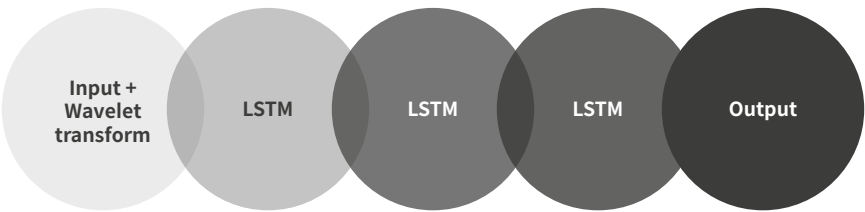
## 概述

人工智能的研发对于我们整个项目当中最为核心且重要，因此我公司正在运营人工智能开发研究室（AI Research Lab）。起初，对于开发人工智能预测模型的时候，我们所采用的方式是深度学习（Deep Learning）。而且，以此为基础，我们也不断将领域扩大到机器学习（Machine Learning），且与深度学习进行整合与统一，并开发出新的模型。

对于我们以往所采纳的深度学习方式进行一个简单的概述，其主要内容是利用循环神经网络（RNN）与卷积神经网络（CNN）方法论来搭建模型系统。为了将该时间序列数据能够投入到模型，其预处理过程很重要。在RNN所投入的数据预处理重点是有效地去除噪声（Noise）。而且，对于投入到CNN的数据预处理重点有效地显示出模式（pattern）。我们不仅单一地采用了对RNN与CNN两种方式，同时也对这两种方式采用了相互互补的方式来尝试搭建该模型。对于在机器学习当中，为了能够完善决策树（Decision Tree）的不稳定性，通常使用到的方法之一，就是采用梯度提升（Gradient Boosting）的方法来搭建相关的预测模型。

我们将从2017年7月开始到至今为止的约有100,000个加密货币时间序列数据用于训练机器学习模型。对于时间序列数据，我们以最高价、最低价、开盘价、收盘价、交易量等五个基本特征为基础来进行了对其数据的分析研究。近期，有关对研究的不断深化，上述的五个基本特征以外，我们还使用了以技术指标为主的特征工程（Feature Engineering）等多个种类的特征点来进行了研发。

有关机器学习模型，对其带入的数据越多，其准确率也会显示出更高。但是，每一个模型都有属于自己的特点。而且，对于时间序列数据的特征而言，其不稳定因素也是不可避免。因此，我们得出的结论是，一味的带入大量的数据，使模型进行学习，未必能够达到预期的效果。实际上，每个模型用不同量数据学习功能将会提高其性能。例如，对于一些CNN模型，在近期以18,000个时间序列数据使让它进行学习时，其模型所列举出的分类准确率是最高的。而且，其中几个长短期记忆网络（LSTM）类型的模型，以全部的数据来使进行学习时，其显示的结果也是达到了最高水平。



[图] Staked wavelet LSTM 模型结构



我们所使用的模型结构如下。如以RNN的方法来进行搭建模型，多层LSTM（长短期记忆网络）或以多层GRU（门控循环单位）而堆叠LSTM 或堆叠GRU（Stacked LSTM/GRU）的功能更佳。而且，通过小波转换的RNN类的模型也显示出了较为满意的结果。我们发现对于CNN模型，卷积层（Convolutional layer）的数量与激活函数（Activation Function）良好的配合模式将能够提高模型功能。层数越多，有可能出现过拟合（overfitting）现象。因此，我们适当地安排Dropout层与正则化。对于LSTM细胞与卷积层，并列的方式层叠出了混合型模型，其显示出了较为满意的功能。对于梯度提升，我们所借用的是XGBoost(Extreme Gradient Boosting), LightGBM, Catboost等梯度提升的框架及算法。

对于Boosting系列的机器学习模型，大部分都比较以依赖特征（Feature）的选择及数据的规模。

## 主要技术介绍

### 概述

我公司拥有专门的人工智能研究团队，其团队人数是10名。我公司也通过该研究团队进行相关人工智能技术的所有开发与研究。我们所采取的主要研究方法及流程如下。第一，对英文、中文以及韩国语圈内的主要研究论进行分析后，实地进行测试与重现。我们从2018年8月开始到2018年12月为止，共参考了约有500篇的论文。实地进行测试的模型当中，出现与我们所预期相符合的模型，我们将其重新进行反复地学习，进一步提高性能。对于有参考价值的研究论文，您可通过我公司的官方网站cosmicbc.com的R&D Blog（研究内容博客）进行相关内容的阅览。有关我公司所积累的研究成果，一直以来都已对外开放，以后也将持续开放。2018年12月开始，我们已经将100篇研究内容应用在实际的服务当中。

对于已搭建的模型，我公司通过亚马逊AWS（Amazon web service），Naver Cloud Platform(NCP)，阿里云（Alibaba Cloud）的人工智能云计算服务器（A.I. Cloud computing server）来正在进行研发。首先，大部分已完成的模型用过去的数据进行学习后，再进行2周~3周的活数据（Live data）学习，最后确认模型的准确率百分比。以此作为参考，我们对显示出较高特定数值以上的模型为重点，用多样化的处理方法，对其模型进行1个月以上的学习。

### RNN

对于循环神经网络（RNN）模型的最大特点是节点定向连接的人工神经网络。与卷积神经网络（CNN）进行对比，该机器学习RNN还拥有CNN不具备的记忆。它的特征是可以有效解决序列数据的预测分类问题的解答。所谓序列数据是，按照顺序排列的组合来构成的时间序列数据，通常是以语音、视频或文本为主。可通过以下序列来表示。

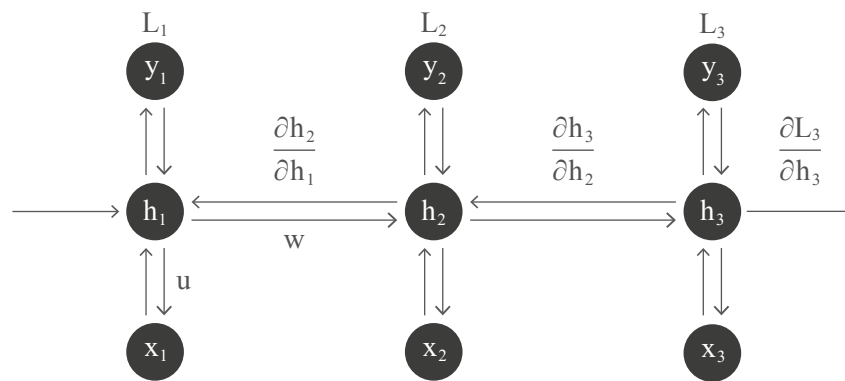
$$x^1, x^2, x^3, \dots, x^T$$

在这里，对于时间序列的长度 $t=1,2,3,\dots,T$ ，通常可以看作可变的序列。以序列时间数据举例，就可被推断出来。有一个文章有第 $t$ 位单词，我们就可以推断出第 $t+1$ 位单词。各个单词是受到前一位单词的影响。RNN抓住对于单词与单词之间的依赖关系，特别是能够对文章的逻辑进行学习后，可预测出较高的准确率。例如，像



加密货币一样的时间序列数据，大部分的情况下现存在的数据都是很容易受到之前数据的影响。同样，RNN模型是拥有像上面所叙述的记忆结构。因此，该结构模型可以临时记住相关信息，并以此信息来表现出不同的反馈。之所以可以把RNN适用于解决加密货币价格的预测分类问题，是因为模型可以根据时间序列数据来获取文章内容的逻辑性，从而进行分类。

进行RNN学习时，需要用到像全连接神经网络一样的梯度下降法 (Gradient Descent)。我们需要对误差进行微分计算，从而计算各层权值。这里可采取两种方法。一是实时递归学习算法RTRL (Realtime Recurrent Learning)，另外则是基于时间的反向传播算法BPTT (Back Propagation Through Time)。前者是对记忆的使用效率较高，后者则是计算速度较快。



[图] BTPP概述

※出自Hands-On Reinforcement Learning with Python by Sudharsan Ravichandiran

BPTT是把RNN以时间方向进行展开，构成类似全连接神经网络一样的结构，之后进行反向传播计算。如同上图所画出的结构图，将每一时刻RNN的整个结构中的环路展开成为时间连续的中间层连接单位。对于中间层的反向传播，是以连续的时间点的中间层单元来进行结合。以这种方式来展开的神经网络，将成为全连接神经网络模样。根据以此来进行反向传播，从而通过计算梯度而可进行反向传播计算。

$$\delta_2^{(3)} = \frac{\partial L_3}{\partial z_3} = \frac{\partial L_2}{\partial h_2} \frac{\partial h_2}{\partial s_2} \frac{\partial s_2}{\partial z_2}$$

$$z_2 = Ux_2 + Ws_1$$

前一层的 $\delta$ 如下。在前一个时间节点的公式如下。

$$\delta_1^{(3)} = \frac{\partial L_3}{\partial z_1} = \frac{\partial L_3}{\partial h_2} \frac{\partial h_2}{\partial s_1} \frac{\partial s_1}{\partial z_1} = \delta_2^{(3)} \frac{\partial h_2}{\partial s_1} \frac{\partial s_1}{\partial z_1}$$

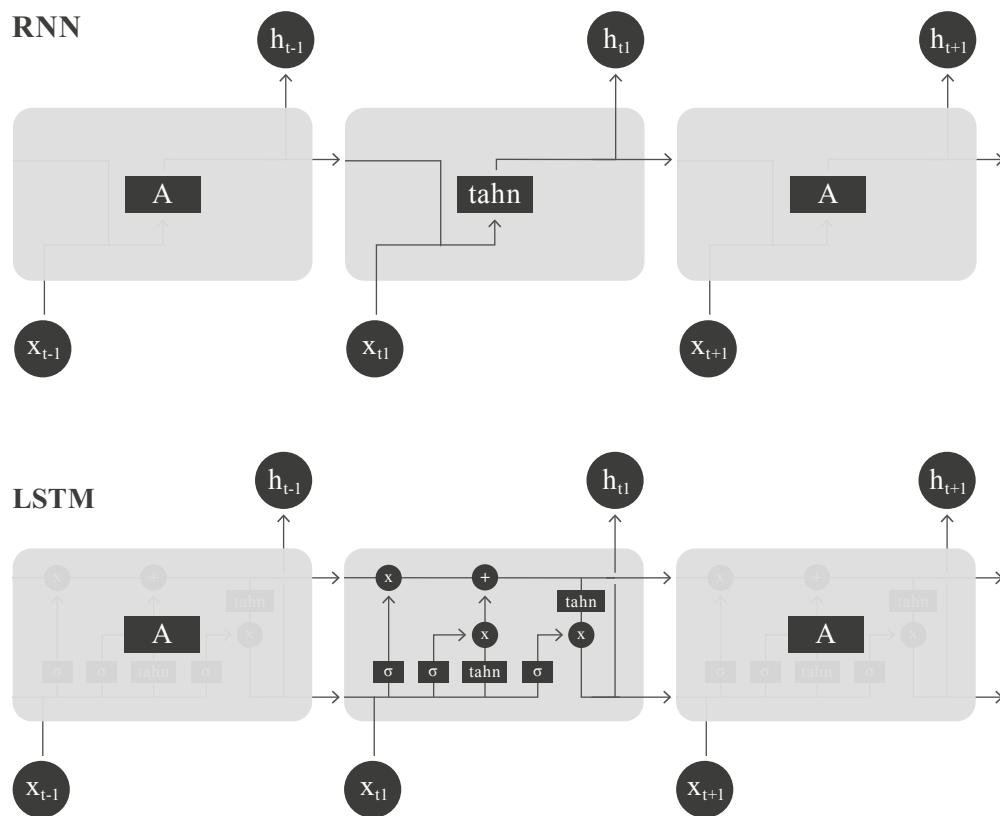
根据上述的内容，RNN可以抓住时间数据的内容，可以进行推断。但是，对于文章的长度很敏感。

理论上，对时间节点t开始至以前的所有时间步骤的信息均可保存。但是，对于长时间的信息是无法进行学习。RNN能够反映出的输出时间最长也就10小时。

这种局限性是因为，对于层数较多的神经网络，通过反向传播计算梯度时，随着经历的层数增多，梯度值很容易发散并消失。所以，RNN很难应对长序列的问题。换句话说，对于短期性的记忆是可以实现，但是相反对长期性的记忆是无法得到实现。

## LSTM and GRU

1990年代初，Sepp Hochreiter, Jürgen Schmidhuber, Youshua Bengio等为了能够解决对于上述内容当中所发生的问题，通过理论性考察而推出了各种方法。其中，最为成功的方法就是长短期记忆网络 (Long Short Term Memory; LSTM)。LSTM与最基本的RNN相比，中间层的各个单元是由记忆单元组成。除此以外的所有的结构都与之前的RNN完全相同。



[图] RNN, LSTM结构的比较

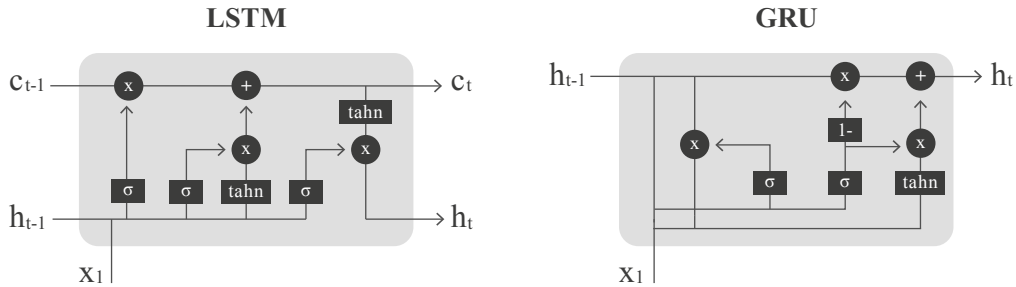
※出自<https://ratsgo.github.io/natural%20language%20processing/2017/03/09/rnnlstm/>

对于LSTM将单元连接理解为传送带 (Conveyotr Belt) 的形式更容易。简单来说，是添加了把信息装载到各个时间步骤来进行传输的方法。在时间序列的某一个时间点上，对于已输出的信息将上传到传送带，并且在必要的时间点的时间步骤下再移动至下来。把重要的信息进行存储，从而防止对那些已经过了一段时间的信号有可能会逐渐消失的问题。

有关对LSTM细胞，以公式来表示如下。在这里， $\odot$ 是表示乘机-哈达马积 (Hadamard product) 矩阵乘法。

$$\begin{aligned}f_t &= \sigma(W_{xh_f}x_t + W_{hh_f}h_{t-1} + b_{h_f}) \\i_t &= \sigma(W_{xh_i}x_t + W_{hh_i}h_{t-1} + b_{h_i}) \\o_t &= \sigma(W_{xh_o}x_t + W_{hh_o}h_{t-1} + b_{h_o}) \\g_t &= \tanh(W_{xh_g}x_t + W_{hh_g}h_{t-1} + b_{h_g}) \\c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\h_t &= o_t \odot \tanh(c_t)\end{aligned}$$

Forget gate  $f_t$ 是会把过去的信息给遗忘的一个门 (Gate)。接收并输入 $x_t$ 与 $h_{t-1}$ 以后，以Sigmoid函数处理后得到的即为输出。如果该值是0，这将会丧失对于以前状态的信息，如果该值是1，它将会完整的记住该信息。“Input gate  $i_t \odot g_t$ ”是为了能够记住当前信息的一个门。接收并输入 $x_t$ 与 $h_{t-1}$ ，Sigmoid函数处理，在以同样的值双曲正切函数处理 (Hyperbolic tangent) 后，两者做乘机-哈达马积 (Hadamard product) 运算后得出的值为Input gate输出值。对于LSTM，已解决了RNN带有的对于长时间记忆的相关问题，在诸多领域中已成功地被使用。LSTM揭示了良好的性能，之后在学界中出现了很多LSTM变种的研究论文。其中，最值得关注的就是由韩国的曹京铉在2014年发表的Gated Recurrent Unit(GRU)。



[图] GRU概述

※出自<http://sqlml.azurewebsites.net/2017/08/12/recurrent-neural-network/>

在上图中，我们可以看到GRU的结构比LSTM的结构更为简单。在细节上，GRU也与LSRM一样是通过门来调整信息的数量，对于这一点两者是相同的。但是，特别是对门的控制方式会有一些不同。

公式表示如下。

$$\begin{aligned}z_t &= \sigma(x_t U^z + h_{t-1} W^z) \\r_t &= \sigma(x_t U^r + h_{t-1} W^r) \\\tilde{h}_t &= \tanh(x_t U^h + (r_t * h_{t-1}) W^h) \\h_t &= (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t\end{aligned}$$

Reset gate  $r_t$  是0~1之间的值, 越接近0就会忘记过去的信息。相反如果是1, 就会记住过去的信息。最后的式子即为改进之处。  $h_{t-1}$  是过去的信息, 而  $\tilde{h}_t$  则是现在的信息, 两者之间的权衡由update gate的  $z_t$  决定。

## CNN

CNN主要是适用于图像认知方面所运用到的神经网络模型算法。其算法是成功地解决了计算机中一直以来的大难题(对于物体的分类认知等)。对于以往的全连接神经网络所面临的问题是, 即使识别较小的字体而计算权值与偏向的个数会暴发性增加。例如, 为了能够辨别出16x16大小的字体, 我们可以设想一下在利用一个拥有100个神经元的隐含层的全连接神经网络。这个网络需要的权值和偏向的数量达到28,326个。如果在该神经网络在增加新的隐含层, 对于需要学习的参数数量将会变得更多。而且, 对于字体的图像如发生竖直或水平的移动, 回转, 变形等变化, 则需要重新输入新的学习数据。

为了能够解决此问题而诞生的CNN是与全连接神经网络是完全不同的。该算法是把层和层之间特定的单元结合的特殊层。该特殊层进行卷积(Convolution)与池化(Pooling)两种对于图像处理相关的计算。

CNN的诞生, 是通过动物大脑的视觉皮质(Visual cortex)的相关神经科学中获得了灵感。更准确来说, 是对于视觉皮质的神经细胞感受野(Receptive field)的局部性与简单细胞与复杂细胞中获得了灵感。同样一个神经, 对于在水平图像产生激活, 或者是在不同角度的图像产生激活所表现出的模样是不一样的。而且, 其他神经也拥有较宽阔的感官领域, 因此会对低水准的模式(Pattern)之间相结合的较为复杂的模式产生激活。通过该规律, 塑造了以下卷积神经网络。如LeNet, AlexNet, VGG, GoogLeNet等都是最具有代表性的CNN模型。这些模型均在图像分类方面展现了比人类高出一等的性能。

CNN的基本构造如下。对于输入图像的第一个卷积层的神经不是连接到已输入的整个图像的像素当中。而是, 只连接在卷积层神经网络的感官领域内的像素当中。但是, 卷积通过线性运算。因此, 为了使模型能够学习更加复杂的非线性关系, 把代表性的非线性激活函数Relu(线性整流函数; Rectified Linear Unit)连接在卷积层内。之后, 为了降低对于池化层的运算量与内存使用量, 以及减少参数的使用数量, 对输入图像做类似下采样处理。根据不同的数据特征, 将卷积层+激活函数+池化层进行不同的组合, 来进行层的堆叠。通常, CNN的输出层连接一层以上的连接层, 有助于高维度图像最后一层做mapping图像分类。

综合以上内容, 在CNN中最为重要的构成因素是卷积层。如有WXW大小的黑白图像, 各个像素标注为(i,j)。然后, 这里还有一个HXH大小过滤器的像素标注是(p,q),  $p, q=0, \dots, H-1$ 。在这里,  $H < W$ 。严格来说, 卷积层的表示公式如下。

$$u_{ij} = \sum_{p=0}^{H-1} \sum_{q=0}^{H-1} x_{i-p, j-q} h_{pq}$$

在卷积层之后的池化层计算如下。假定, 在大小为WXWXK图像中, 像素(i,j)点为中心有一个大小为HXH正四边形的框, 做Padding计算处理图像边缘都包括的  $P_{ij}$  可以求出来。对于该  $P_{ij}$  内的像素, 每一频道K都会单独地利用  $H^2$  个的像素值来获得一个像素值  $u_{ijk}$ 。代表性的求这个像素值的方法为最大池化(Max Pooling), 平均池化等。

最大池化是在 $H^2$ 个的像素值当中选择最大值的方式，该计算方法如下。

$$u_{ijk} = \max_{(p,q) \in P_{ij}} z_{pqk}$$

平均池化是获取对 $H^2$ 个的像素值的平均值，以此作为像素值。

$$u_{ijk} = \frac{1}{H^2} \sum_{(p,q) \in P_{ij}} z_{pqk}$$

最后，包括了最大池化与平均池化的 $L_p$ 池化。该池化运算是在各自的频道上并行计算。换句话说，池化层的输出频道数与输入图像的频道数是一样的。 $L_p$ 池化的公式表示如下。

$$u_{ijk} = \left( \frac{1}{H^2} \sum_{(p,q) \in P_{ij}} z_{pqk}^p \right)^{\frac{1}{p}}$$

我们从该公式中不难看出， $P=1$ 就是平均池化， $P=\infty$ 就是最大池化。

对于价格预测分类问题中，我们可以将时间序列数据的变化模式化。我们得出结论，根据卷积神经网络 (CNN) 而搭建的模型是拥有强大的预测能力。

## 梯度提升 (Gradient Boosting)

有关梯度提升是对于梯度下降(Gradient Descent)与Boosting相结合的合成词。Boosting是相对来说只是把单一决策树进行构建，一步步进行学习。基于此结果再建一个决策树。梯度提升方法是适用于所有损失函数可微分的函数。

梯度提升算法如下。

首先对学习机(h)进行初始化：

$$h_0(x) = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, \gamma)$$

对于，各个 $m=1,2,\dots,M$ 可以1,2,3,4来进行反复。

$$r_{im} = - \left[ \frac{\partial L(y_i, h(x_i))}{\partial h(x_i)} \right]_{h(x)=h_{m-1}(x)},$$

$i = 1, 2, \dots, m$

1. 计算伪残差(pseudo-residual)。

$$\{(x_i, r_{im}) \mid i = 1, 2, \dots, n\}$$

2. 计算基本学习机的 $h_m(x)$ 的伪残差。  
训练的数据为

$$h_m(x) = h_{m-1}(x) + \gamma_m h_m(x)$$

4. 更新学习机。

$$\gamma_m = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, h_{m-1}(x_i) + \gamma h_m(x_i))$$

3. 计算乘数 $\gamma_m$

$$h_M(x) = \sum_{i=1}^M \gamma_i h_i(x) + h_0(x)$$

最终选定的模型如下。

对于价格变动的分类问题，如果使用的指数(Exponent)的损失函数，其公式为 $r_{im} = y_i \exp(-y_i h_{m-1}(x_i)) = y_i w_{im}$ 。如 $w_{im}$ 的值是正分类，其值为 $e^{-1}$ 。相反，误分类的话，其值为 $e^1$ 。也就是说，下一阶段会给误分类更大的权值。所谓的伪残差是一种残差。对于Boosting的一般化方法是根据损失函数被定义之后，以最梯度发来对损失函数降低到最低的过程。

以梯度提升方法而搭建的模型中,最具有代表性的就是XGBoost, LightGBM, Catboost等。

首先是XGBoost。以并列处理的方式,能够快速地进行学习与预测。能够提供多样化的定制型选项内容。根据Greedy算法来自动剪枝防止过拟合(Overfitting)。对于其他算法之间连接性较好。

与XGBoost相比, LightGBM的最大优势就是学习时间很短,内存使用量也相对较少。LightGBM是在XGBoost出现两年后才开发出来的。该模型是继承了XGBoost的优势,也弥补了诸多缺陷。LightGBM是采用了叶明智策略分裂方式,不考虑叶子的均衡,在损失值的最大的叶子节点上持续性的分裂,从而发展下去,所以会出现树的非对称现象。但是,相对于均衡树分割方法,可将误差损失最小化。

最后是Catboost。在该模型所引入的Ordering Principle概念解决了对以往的因目标泄露(target or data leakage)而导致的位移预测现象(prediction shift)与分类变量(Categorical Variable)的预处理问题。换句话说,在计算leaf value的过程中,所采用到的方法就是Ordered Boosting。以往的梯度提升是在搭建新树的时候,将根据现有的模型中所使用到的数据进行再利用,因此存在着过拟合。该模型与以往的模型所不同。该过程是先求出逆向的leaf value,之后定出树结构的。



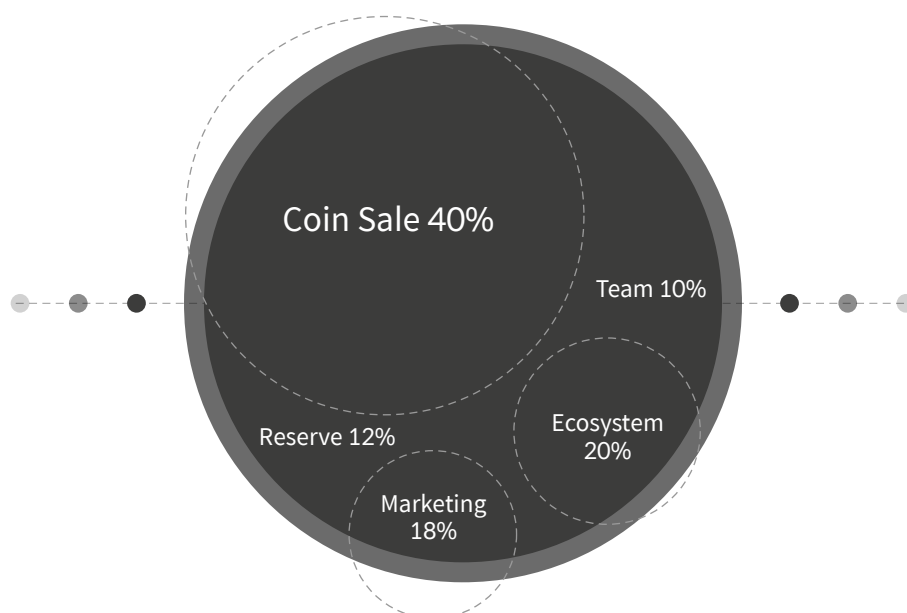
# 通证经济体系

## 加密货币介绍

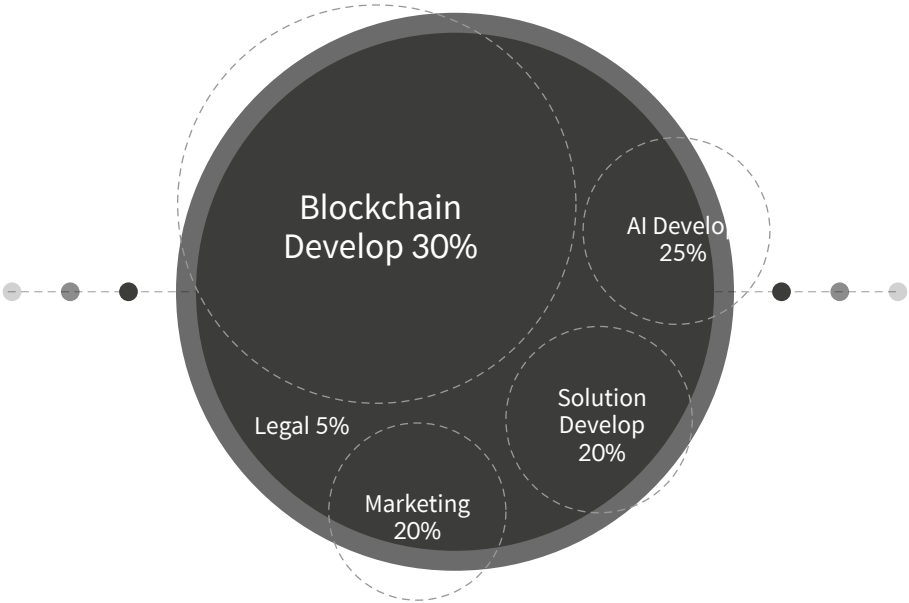
总发行量	12亿枚
单位价格	1AIPE = 约 0.02 USD
名称	AI Prediction & Exchange Solution
Coin Symbol	AIPE
标准	ERC-20
Decimal类型	类型: 18
Hard Cap	412,500,000 AIPE
Soft Cap	137,500,000 AIPE

## 分配内容

AIPE币是针对以平台的开发与AI预测为基础，搭建相关生态体系为目的而进行发行。对于参加ICO及IEO的用户，可获得整体发行量的40%的流通数量。除了加密货币销售 (Coin Sale) 之外的剩余货物量的30%，是用于构建初期的生态体系。例如，AI预测平台的搭建，营销等方面。对于APIE Reserve部分的15%的货物量，将会使用在对于扩展平台以及后期开发而需要的基础设施的搭建等。Reserve货物量的15%与团队的分配货物量的15%将按照顺序而进行具体的分配。

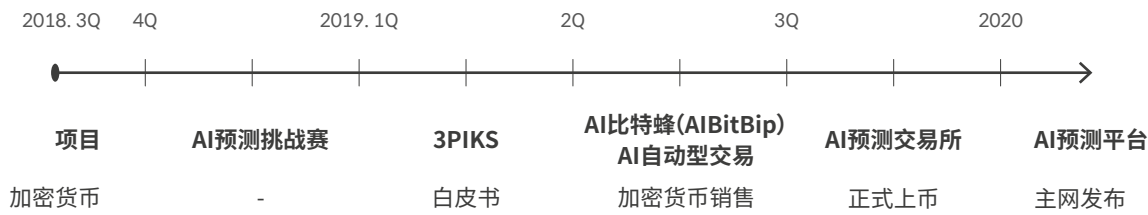


款项运营



服务项目	具体内容
区块链研发	根据区块链的研发,我们将保持AI研发平台。同时,对于从中所开发出的各项模型,我们也将通过该研发平台进行对权利的保护。 作为一个有公信力的平台,我们将大力支持把AI预测数据库内容储存到区块内的相关研究。
人工智能研发	对于AI研究人员的教育及研究方面构建优良环境,大力支持预测模型的研发。 坚持以机器学习为基础的大数据分析算法的相关研究。
解决方案研发	能够提供综合解决方案的相关项目的开发。 3PIKS, AI比特蜂 (AIBitBip), AI自动型交易模式, 交易所, 保证金交易所等项目的开发。
营销	现区块链的产业领域的竞争异常激烈。根据对此现状的考虑,为了能够使我公司的平台迅速地适用到实际的商业模式当中,我们将积极地利用各种营销渠道,且能够确保使用者人数。
法律业务费	对于服务项目的开发及运营当中所发生的各种版权及相关法律方面的处理费用。

路线图



## 通知及注意事项

我们在此郑重提示，AIPX项目白皮书是以够详细地地介绍我公司的整体项目，商业模式以及相关技术等内容为目标，并不是以诱导投资为目的。对于阅览本白皮书的所有人群，凡是根据白皮书上所记载的内容作为参考而进行投资，以此发生的任何损失、损害、债务等各项财务方面的问题及发生都与我们的AIPE团队无关，我们也对此不承担任何的责任。

而且，对于阅览该白皮书的人，在进行自我判断与决定时，通过使用本白皮书（包括，对于该白皮书进行参考，以白皮书作为依据，但也不仅限于此内容。）而发生的所有财政以及债务等方面所发生的损失，我们AIPE团队一概不负任何的赔偿以及补偿等相关责任。

AIPE项目白皮书从“编写完毕的时间点开始”到“未来的任何一个时间点”为止，所发生的任何内容相关的变更或内容是否属实等事项，我们对此一概不负责任。关于本白皮书，根据AIPE团相关政策或者是公司内部决定，随时都可能会进行对内容上的更改。因此，AIPE团队与本白皮书有连系的，对于阅览此白皮书的相关人群，以及各项内容的陈述，均不承担任何的法律责任。

首先，AIPE白皮书内容是符合法律及规章制度。而且，对于该白皮书，是否有侵害到第三方的权利，或者是否在商业领域有一定的价值，或者是阅览该白皮书的人群是否可以达到其预想的效果，白皮书的内容是否完全属实等，我们也对此不承担任何的责任。不仅如此，对于免责相关的内容不仅限于以上内容。

如没有获得AIPE团队的同意，本白皮书不允许随意复制，使用或向外界公开。

