

AI Prediction & Ecosystem White Paper ^{KR}

AIPE COIN: TABLE OF CONTENTS

Abstract	01
Services	02
3PIKS	02
AIBitBip	02
AI Auto-Trading	02
AI Prediction Platform	03
Exchange Solution	03
AI Prediction Ecosystem	04
AI Technology	05
Outline	05
Introducing Main Technology	06
TOKEN Economy	14
Coin Structure	14
Allocation	14
Ecosystem	15
Road Map	15
Notice & Caution	16

ABSTRACT

‘인공지능(Artificial Intelligence; 이하 AI)’은 이세돌 9단과 알파고의 대국으로 세상에 널리 알려짐과 동시에 큰 반향을 가져왔고, 이후 전 세계적으로 다양한 분야에 이용되기 시작하였다. 금융과 관련된 서비스 및 보험(BF-SI)분야를 비롯하여 생명과학과 교육 연구분야에서 활발히 이용되고 있고 정부, 국방, 제조, 에너지, 운송, 의료를 포함한 각종 산업에도 이용될 뿐만 아니라 소매, 통신, 미디어, 엔터테인먼트, 스포츠에 이르는 생활부분까지 넓게 이용되고 있다.

실제로 지능형 비즈니스 프로세스에 대한 요구가 증가되면서 사용자에게 전달되는 전반적인 경험이 향상되고 있다. 이에 발맞추어 Google, Microsoft, IBM, AWS, Intel, FICO(Fair Isaac Corporation), Salesforce, Baidu, SAS 등의 글로벌 기업들은 AI 기능을 기존 비즈니스 응용 프로그램과 통합하는 데 주력하고 있다.

글로벌 시장조사기관 Research and Markets의 발표에 따르면 AI 서비스 시장은 2018년 7월 기준 다가오는 2023년까지 연평균 성장률이 48%이상 예상되며, 이 성장 기반에는 지능형 비즈니스 프로세스에 대한 필요성 증대와 IoT, Connected Device의 증가로 인한 대용량 데이터 생성과 부가되는 운영비용 감소에 대한 필요성 등이 작용했다고 분석했다. 더불어 기존 인프라 및 스마트시티 프로그램을 향상시키는 세계 각 정부와 기업, 지자체의 노력이 증가하면서 전반적인 AI Service Market 성장을 촉진하고 있으며 이는 지속될 것으로 전망하고 있다.

이처럼 AI 기술은 지속적인 성장을 할 수 있는 높은 부가가치의 산업이고 실제 수익 창출이 일어나는 시장임을 인지해야 한다. 하지만 외형적 성장과 그 가능성에 반해 실제 AI 기술이 시장과 접한 문제와 그에 대한 해결책은 아직 결음마 수준인 것이 현실이다.

이에 우리 AIPE는 AI Prediction을 중점으로 Prediction Model을 구축하고 이 데이터를 테스트-상용화에 접목하길 원하는 AI 연구자와, AI Model 데이터를 암호화해 투자에 지표로써 활용하길 바라는 투자자, 그리고 축적된 AI Developer pool과 AI 데이터를 필요로 하는 기업 모두를 상호 연결해주는 블록체인 기반의 플랫폼을 개발중이며 연구자, 투자자, 기업이 느끼는 갈증을 해결하는 것이 목표이다.

본 플랫폼 내에서 생산되고 공유되는 모든 정보는 블록체인 기술을 통해 안전하게 관리되고, 모든 AI 연구자와 AI Model을 사용하는 투자자들은 자신이 창출한 콘텐츠 및 개인 정보에 대한 소유권과 관리 권한을 행사할 수 있으며, 이를 기업 참여자에게 공유하여 기여하는 경우 이에 대한 보상을 받게 된다.

또한, 모든 참여자와 기업들은 AI Prediction Platform이 제공하는 API를 이용하여, 본 플랫폼 내에서 다양한 On Demand서비스를 구현할 수 있다. 우리 AIPE팀은 AI Prediction Platform을 통해 AI 연구자가 보유한 Prediction Model과 실제 사용자들의 데이터소유권 및 관리권한을 필요한 곳에 재분배하고 이에 대해 명확히 보상함으로써 필요한 가치들이 선순환 되도록 하고자 하며, 이를 기반으로 AI 산업의 혁신에 기여하려는 목표를 가지고 있다.

SERVICES

3PIKS

AI와 블록체인기술을 접목한 독창적인 리워드 플랫폼

우리는 ‘인공지능(AI)’ 과 ‘인간의 판단(Human Factor)’이 결합된 암호화폐 가격예측 서비스를 제공하고자 AI와 암호화폐 시장을 접목한 3PIKS를 개발했다.

암호화폐 시장 참여자들은 3PIKS 서비스 내 AI Model의 가격예측 뿐만 아니라 3PIKS 사용자들의 예측이 모여 형성된 결과를 집단지성의 한 형태로써 참고할 수 있다. 유저들은 자사에서 자체적으로 개발한 AI Model과 함께 암호화폐의 가격을 예측하고 그에 대한 보상으로 AI Token을 받게 된다. AI Token은 추후 AIPE Coin으로 교환할 수 있고 전략적 파트너십을 맺은 파트너사들의 토큰 및 코인들과 교환될 예정이다. 이와 같이 사용처에 대한 다양성을 가지는 방법으로 생태계 구성을 한다.

AIBitBip

AI 기반 암호화폐 가격변동 예측신호 서비스

AI Model이 매매에 도움을 줄 수 있도록 의미 있는 상승/하락 징후를 포착하여 이를 신호의 형태로 알림 서비스화 한다. 이용자들의 접근성을 높이기 위해 어플리케이션의 형태로 프로토타입이 개발되었으며 2019년 2분기 런칭할 예정이다. 초기에는 자체적으로 개발한 수 십여 종의 Model을 우선적으로 서비스하고, 추후 AI Prediction Platform에서 개발된 여러 연구자들의 Model을 선별하여 추가 및 확장하여 선보일 예정이다.

AIBitBip의 가치는 단순히 매매에 대한 정보를 알려주는 것에 그치지 않는다. 불특정 다수가 해당 신호에 맞춰 매매를 하게 된다면, 이는 하나의 큰 집단으로 분류할 수 있다. 동일한 시점에 많은 양의 자본이 투입되어 가격의 흐름을 만들게 된다면 투자에 대한 리스크를 대폭 낮출 수 있을 것으로 예상되고, 이용자는 이를 매매에 대한 참고 자료로 이용 할 수 있을 뿐만 아니라 매매에 대한 리스크 또한 줄일 수 있을 것으로 전망한다. 어플리케이션의 특성상 편리하고 접근성이 용이하여 이용자들에게 노출도가 높다. 이는 어플리케이션 내 각종 광고 게재 및 Referral 시스템 사용처로 확장이 가능하다.

AI Auto-Trading

AI Prediction Signal을 바탕으로 한 Auto-Trading Bot

AIBitBip을 통해 Signal을 받는다고 해서 누구나 매매로 수익을 볼 수 있는 것은 아니다. 24시간동안 대규모로

거래가 이루어지는 암호화폐 시장에서 사용자가 Signal을 받고 실시간으로 매수, 매도를 하는 것은 어려울 것이다. 이를 해결하기 위해 우리는 각 거래소 API를 이용한 AI Auto-Trading Bot을 서비스할 예정이다. 해당 프로젝트를 성공적으로 이끌기 위해 오랜 시간 AI Prediction Signal을 통해 매매를 시도해왔다. 많은 시행착오가 있었지만 AI의 학습이 거듭될수록 적중률과 수익률의 점진적인 상승을 볼 수 있었다..

정확한 타이밍에 매매를 하여 수익을 볼 수 있는 서비스를 제공하여 이용자가 수익을 얻는 것에 초점을 맞추고 있다. 해당 서비스를 이용하기 위해서는 AIPE를 일정이상 보유하고 있어야 하며 Auto-Trading의 수익 일부를 연구자들과 공유함으로써 생태계를 유지한다.

또한 Auto Trading은 단순히 수익만 창출해내는 것이 아니라 Exchange와도 긴밀한 협력관계에 놓일 수 있다. 해당 Auto-Trading을 진행하기 위해서는 수많은 호가등록과 체결의 과정을 반복해야 한다. 이를 통해 Exchange는 많은 거래량과 화폐 유동성을 확보 할 수 있어 Market을 형성하는데 중요한 기능을 한다. 우리는 B2B 비즈니스를 통해 이용자와 Exchange, 그리고 Prediction Platform 발전에 상호 도움이 되는 생태계를 형성할 수 있다.

AI Prediction Platform

AI 연구-개발 환경 개선과 검증된 Data확보를 해결할 수 있는 AI Prediction Platform

현재 AI 연구에 있어 가장 큰 진입 장벽은 다량의 데이터 정제하는 것과 Computing Power의 확보일 것이다. 개인이 이러한 환경을 세팅하는 것은 불가능에 가깝고, AI Model 데이터를 필요로 하는 기업의 상황 또한 별반 다르지 않을 정도로 환경은 열악하다.

AIPE에서는 이러한 AI 연구개발 환경을 개선하기 위하여 AI Prediction Platform을 도입하고자 한다. Platform에서 AI 연구자들은 Open Source Library 환경에 손쉽게 접근-이용할 수 있고, 기업은 Traffic과 Model 데이터를 확보할 수 있다. 이러한 환경에서, 많은 AI 연구자들이 Platform에서 제공하는 데이터 툴을 사용하여 다양한 AI Model을 만들어낸다고 하면, 클라우드 제공기업의 시장 진입장벽이 낮아지게 된다. 이를 통해 체계적인 AI Model의 관리와 데이터의 공유가 수월해지고, 서버 사용비용 감소, 다양한 트레이딩봇/알고리즘의 증가, 사용자 접근성 또한 증가함으로써 새로운 자산운영 트렌드의 시작점이 될 수 있다고 판단된다.

이렇듯이, AIPE가 제공하는 Platform Infra는 AI 연구자와 투자자, 그리고 기업의 편의에만 한정되는것이 아닌, 인공지능 산업 전반의 혁신을 가져온다는 점에서 큰 의미가 있다.

Exchange Solution

많은 기업들이 암호화폐 거래소 개발과 서비스에 도전하지만 시간적, 물리적 비용과 더불어 기술적인 전문성 부족, 마케팅 부재 등으로 난항을 겪고 있다.

AIPE는 이러한 문제를 해결할 수 있도록 자체적으로 개발한 암호화폐 거래소의 솔루션을 제공하려 한다. 기존 중앙화 거래소의 솔루션을 제공함은 물론, 마진 거래와 레버리지 거래를 지원하는 암호화폐 마진 거래소를 개발 중에 있다. 당사가 개발중인 마진 거래소는 ‘BitMEX’와 유사한 마진 선물거래를 지원하고, 추가로 마진 옵션거래까지 지원할 예정이다.

많은 암호화폐 거래소 솔루션 업체가 존재하지만 해당 업체들은 중앙화 거래소의 제공에 국한되어 있다. AIPE는 마진 거래가 가능한 거래소뿐만 아니라 파생상품인 선물거래, 옵션거래 제공에 마진거래를 더한 마진 선물거래, 마진 옵션거래까지 지원하는 암호화폐 투자솔루션을 제공하려 한다.

AI Prediction Ecosystem

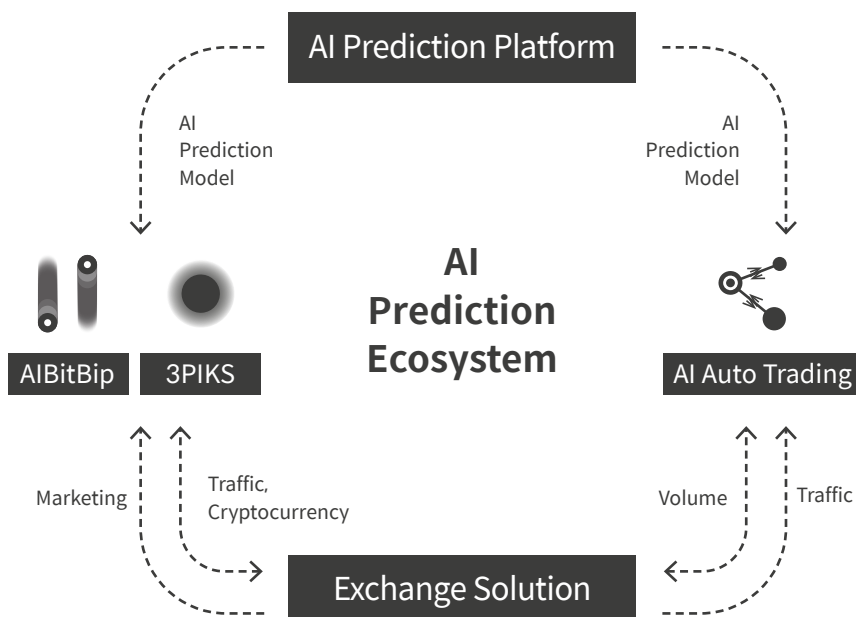
AIPE의 AI 서비스들이 공존하는 Ecosystem 비전

AIPE가 추구하는 비전은 각 서비스의 고유가치로 하나의 유기적 암호화폐 생태계를 이루는 것이다.

리워드 플랫폼인 3PIKS와 암호화폐 가격변동 알림 서비스인 AIBitBip, 두 서비스를 이용하여 사용자들에게 코인 및 토큰을 Airdrop할 수 있고 거래소나 특정 코인의 마케팅 채널로 활용이 가능하다.

또한, 3PIKS에서는 AI Token(가칭)을 보상받아 협약사의 Token 및 Coin으로 Swap가능하도록 진행중이다. Swap된 코인은 자사의 마케팅 서비스를 지원하는 거래소에서 출금 가능하고, 이를 통해 거래소의 Traffic과 유동성을 확보할 수 있다.

이처럼, AIPE의 Ecosystem은 단순히 우수한 AI Model의 확보나, 거래소 개발에 필요한 솔루션만을 제공하는 것이 아니다. 각 서비스가 상호 유기적인 관계와 가치를 지니며 암호화폐 가격예측, 화폐 유동성, 홍보, 트래픽 그리고 시장의 안정성을 순환시키는 구조를 가진다.



AI TECHNOLOGY

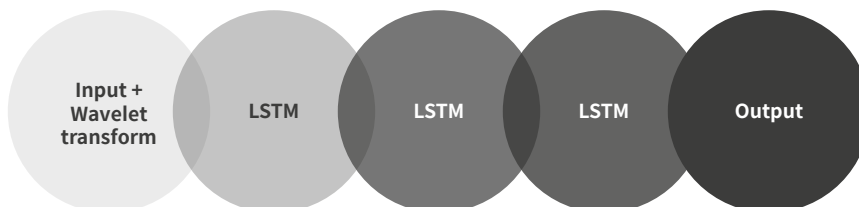
outline

AI 개발이 가장 핵심인 Project이기에 우리는 AI Research Lab를 운영하고 있다. 처음 Prediction Model을 개발할 때는 딥러닝의 방식의 모델을 이용했지만 점차 확대하여 머신러닝과 딥러닝 모두를 통합한 모델들이 개발되어 나오고 있다.

기존 딥러닝의 방식을 개략적으로 표현한다면 RNN과 CNN의 방법론을 이용한 모델링 시스템의 구축이다. 해당 시계열 데이터를 모델링에 투입하기위한 전처리가 매우 중요하다. RNN의 투입되는 데이터 전처리는 Noise를 효과적으로 제거하는데 중점을 두고 있다. 또한 CNN에 투입되는 데이터 전처리는 패턴을 효율적으로 나타내는데 중점을 나타낸다. 이러한 RNN과 CNN의 방식을 독립적으로 이용하기도 하지만 상호보완적으로 모델링을 시도해보았다. 머신러닝 방법론으로는 결정나무(Decision Tree)의 불안정성을 보완하기 위해 일반적으로 사용되는 방법 중의 하나인 그래디언트 부스팅(Gradient Boosting)을 이용하여 예측 모델을 구성하고 있다.

머신러닝 모델들의 훈련을 위해 2017년 7월경부터 지금까지 약 100,000여 개의 암호화폐 시계열 데이터를 사용했다. 시계열 데이터로는 고가, 저가, 시가, 종가, 거래량 등의 5개의 Feature를 바탕으로 이용했다. 최근 연구가 가속화됨에 따라 상기 5개의 기본 Feature이외에도 기술적 지표 등을 이용한 Feature 엔지니어링을 통해 만들어진 다양한 Feature들을 사용한다.

머신러닝 모델은 투입 데이터가 많으면 더 높은 정확도를 성취하는 경향이 있다. 하지만 모델마다 특징이 다르고, 시계열 데이터의 특성상 변동적인 요소들이 있기 때문에 무조건 많은 데이터를 투입해서 모델을 훈련시키는 것이 능사는 아닌 것으로 판단된다. 실제 각각의 모델들은 선별적인 데이터로 훈련시켰을 때 높은 성능을 보였다. 예를 들면, 특정 종류의 CNN 모델은 최근 1만 8천여개의 시계열 데이터로 훈련시켰을 때, 분류 정확도가 가장 높았으며, 몇몇 LSTM류의 모델은 전체 데이터로 훈련시켰을 경우, 가장 높은 성능을 나타냈다.



[그림] Staked wavelet LSTM 모델의 구조

다음으로 우리가 사용하는 모델의 구조는 간략히 다음과 같다. RNN 방법론을 이용하는 경우, 복수의 LSTM 혹은 GRU 층을 쌓은 Stacked LSTM/GRU가 좋은 성능을 보였다. 또 Wavelet 변환을 거친 RNN 류의 모델도 만족할

만한 결과를 보였다. CNN 모델은 합성곱층의 개수와 활성화함수의 조합이 성능을 제고하는데 중요한 것으로 판단되었다. 층이 많을수록 과적합의 문제가 있어, 드롭아웃과 배치 정규화를 적절히 사용했다. LSTM 셀과 합성곱층을 병렬로 쌓은 혼합 모델도 좋은 성능을 보였다. 그래디언트 부스팅의 경우, XGBoost, LightGBM, Catboost 등의 방법론을 차용하였다. 부스팅 계열의 머신러닝 모델의 경우, Feature 선택과 데이터의 규모에 성능이 의존하는 경우를 보인다.

Introducing Main Technology

Outline

당사는 10여명 규모의 A.I. 연구팀에서 해당 연구를 진행중이며, 연구방법은 다음과 같다. 우선 영어, 중국어, 한국어권의 주요 연구 논문들을 연구해서 그 중에서 대부분을 실제로 구현해본다. 그를 통해 2018년 8월부터 2018년 12월까지 약 500여편의 논문을 참고하였으며, 실제 구현한 모델 중에서 당사에 기준에 적합한 성과를 보이는 모델을 다시 지속 적으로 학습시키며, 이용방법을 개선한다. 유의미한 참고 연구 논문에 대해서는 cosmicbc.com의 R&D Blog를 통해서 확인할 수 있으며, 당사는 해당 연구성과를 꾸준히 공개해왔으며, 앞으로도 계속 공개해 나갈 예정이다. 현재 2018년 12월 시점으로 100여 편의 논문의 내용을 실제 서비스에 구현했다.

당사는 그와 같이 구현한 모델을 Amazon web service(AWS), Naver Cloud Platform(NCP), Alibaba Cloud의 A.I. Cloud Computing Server를 이용해 구현해서 연구를 진행중이다. 우선 대부분의 모델이 구현된 상태에서 과거 데이터를 통해 학습을 시키고 2-3주간의 Live 데이터 학습을 통해서 실제로 어느 퍼센트 정도의 정확도를 갖추는지 확인한다. 이를 기준으로 특정 수치 이상의 확률을 보이는 모델을 중심으로 다양한 변형 처리 방식을 구현하여 1개월 이상 학습시킨다.

RNN

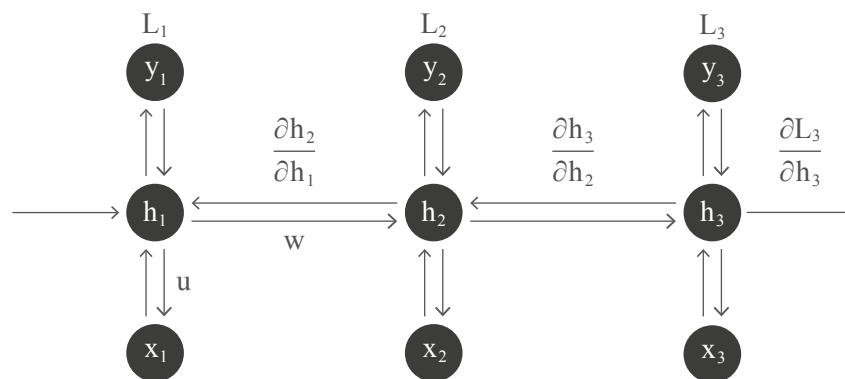
RNN(Recurrent Neural Network) 또는 순환신경망 모델의 특징은 완전 연결 신경망이다. 후술할 CNN(Convolutional Neural Network; 합성곱신경망 또는 컨브넷)에 없는 메모리를 가진다. 이는 연속열 데이터의 예측 분류 문제를 풀기에 적합하다고 알려져 있다. 연속열 데이터란 각각의 요소에 순서가 있는 모임으로 주어지는 시계열 데이터로써 음성, 동영상, 텍스트 등을 예로 들 수 있다. 이는 다음과 같은 형태로 표현이 가능하다.

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

여기서 연속시계열의 길이 $t=1,2,3,\dots,T$ 는 일반적으로 가변적이다. 실제 연속열 데이터를 다루는 추정 문제의 예로, 어떤 문장의 t 번째의 단어까지 주어졌을 때, $t+1$ 번째 단어를 예측해 볼 수 있다. 각 단어는 이전 단어의 연속열에 강하게 영향을 받기도 한다. RNN은 단어 간의 의존관계, 즉 문맥을 잘 학습하여 높은 정확도의 예측을 이뤄내고 있는 것으로 평가된다. 암호화폐 같은 시계열 데이터의 경우 이전 데이터가 현재 데이터가 영향을 주는 경우가 다반사이다. RNN은 상기 서술한 메모리 구조 덕분에 정보를 일시적으로 기억하고 그에 따라 반응을 달리 할 수

있다. 연속적 시계열 데이터의 ‘문맥’을 포착하고, 분류 문제도 잘 처리할 수 있다. 즉, RNN을 암호화폐 움직임 예측 분류 문제를 푸는데 적용시킬 수 있다.

RNN을 학습시킬 때는 완전 연결 신경망과 같이 경사하강법(Gradient Descent)을 사용한다. 각 층의 가중치에 대한 오차의 미분을 계산해야 하는데, 여기에 두 가지 방법을 적용시킬 수 있다. 하나는 Realtime Recurrent Learning (RTRL), 다른 하나는 Back Propagation Through Time (BPTT)이다. 전자는 메모리 사용 효율성이 높은 반면, 후자는 계산 속도가 빠르다.



[그림] Staked wavelet LSTM 모델의 구조
※출처 Hands-On Reinforcement Learning with Python by Sudharsan Ravichandiran

BPTT는 RNN을 시간 방향으로 전개하여 완전 연결 신경망과 같은 구조로 변경한 후, 역전파 계산을 실행한다. 상기 그림과 같이 매 시각의 RNN을 펼쳐 놓아 중간층의 귀환로를 연속되는 시각의 중간층 유닛으로의 결합으로 표현한다. 이렇게 전개한 신경망은 기존의 완전 연결 신경망이 되며, 그에 따라 역전파를 통해 기울기를 계산할 수 있게 된다. 기존의 역전파처럼 이전 층으로 전파되는 델타를 정의할 수 있다.

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

$$\text{여기서 } z_2 = Ux_2 + Ws_1$$

이전 시간 스텝에서도 같은 방법을 적용하면, 아래와 같이 표현할 수 있다.

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

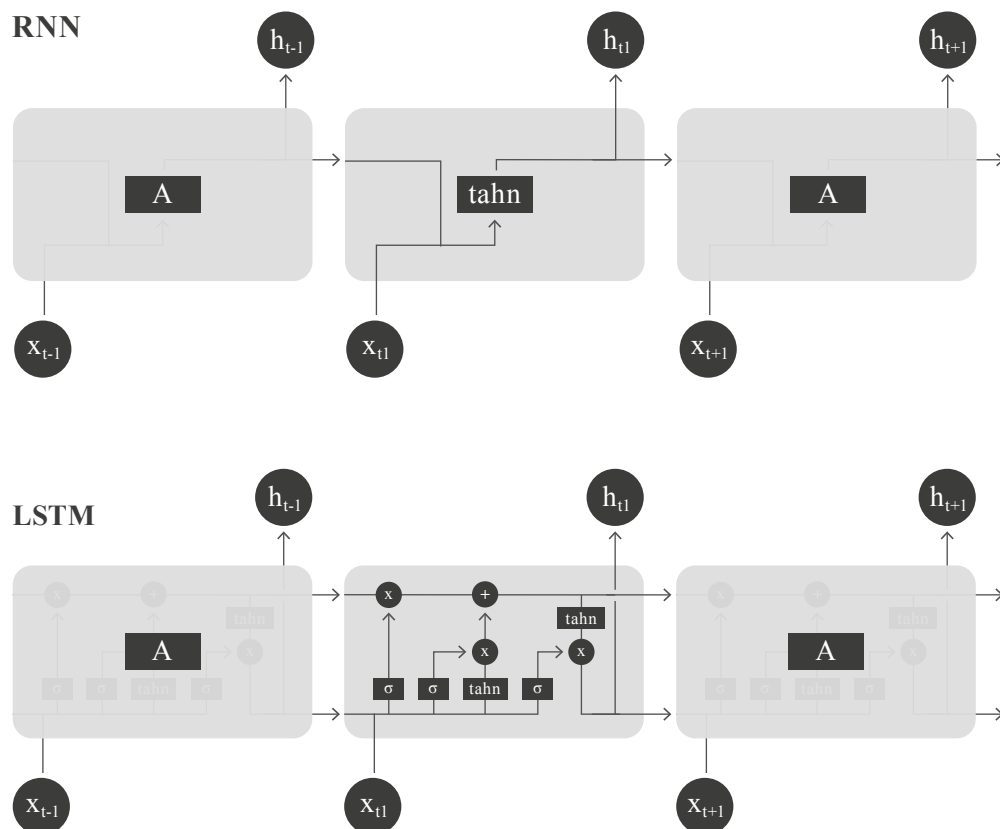
앞에서 서술한 바와 같이 RNN은 연속열 데이터의 문맥을 포착, 추정이 가능하지만 이때 포착할 수 있는 문맥의 길이에 민감하다.

이론으로는 시간 t 부터 이전의 모든 타임스텝의 정보를 유지할 수 있지만 긴 시간에 걸친 의존성은 학습할 수 없는 문제가 발생한다. 실제로 RNN의 출력에 반영시킬 수 있는 것은 최대 과거 10시각 정도로 파악되었다.

이러한 한계는 층수가 많은 신경망에는 역전파법으로 기울기를 계산할 때 층을 거슬러 올라감에 따라 기울기 값이 발산하거나 소멸하기 쉬운 성질이 부여된다. 즉, RNN의 층수가 적다고 해도 역전파 계산 시에는 많은 층수를 다루는 것과 동치가 되며 그에 따라 기울기 값도 발산하거나 소실되기 쉽다. RNN에서 긴 연속열을 다루기 어려운 것은 이 이유 때문이다. 단기적인 기억은 실현할 수 있지만 장기적인 기억을 실현하는 것은 어렵다는 말과 같다.

LSTM and GRU

1990년대 초 호크라이터(Hochreiter), 슈미트후버(Schmidhuber), 벤지오(Bengio)가 앞에서 발생한 문제에 대한 이론적인 고찰을 통해, 이를 해결하기 위한 방법 몇 가지를 소개했다. 그 중에서도 가장 성공적인 방법은 장·단기 기억(Long Short Term Memory; LSTM)으로 알려져 있다. LSTM은 앞서 설명한 기본적인 RNN에 비해서 중간층의 각 유닛이 메모리 유닛이라 불리는 요소로 구성된 구조이다. 그 외의 구조는 기존 RNN과 완전히 동일하다.



[그림] RNN / LSTM 구조 비교

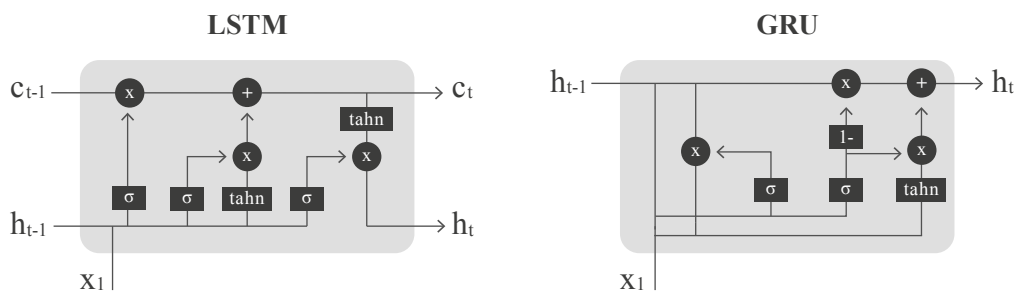
※출처 <https://ratsgo.github.io/natural%20language%20processing/2017/03/09/rnnlstm/>

컨베이어 벨트 역할을 수행하는 셀 상태(Cell-state)를 상상하면 이해하기 쉬워진다. 정보를 여러 타임스텝에 걸쳐 이동시키는 방법이 추가되었다. 시계열의 어느 시점에서 추출된 정보가 컨베이어 벨트로 올라가 필요한 시점의 타임스텝으로 이동하여 내려온다. 나중에 위해 정보를 저장함으로써 처리 과정에서 오래된 시그널이 점차 소실되는 것을 막아주는 것이 LSTM이 하는 일이다.

수식으로 풀어보면, LSTM 셀의 수식은 아래와 같고, 여기서 \odot 는 요소별 곱셈을 뜻하는 Hadamard product 연산자이다.

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

Forget gate f_t 는 과거 정보를 잊게 해주는 게이트이다. x_t 과 h_{t-1} 를 입력 받아 시그모이드를 적용시킨 값이 forget gate가 내보내는 값이다. 그 값이 0이라면 이전 상태의 정보를 잊고, 1이라면 그것을 완벽히 기억한다. Input gate $i_t \odot g_t$ 는 현재정보를 기억하기 위한 게이트이다. x_t 과 h_{t-1} 를 입력받아 시그모이드를 적용하고 또 같은 입력으로 하이퍼볼릭 탄젠트(tanh)를 적용시킨 후 Hadamard product 연산을 한 값이 Input gate가 산출하는 값이다. 요컨대, LSTM은 RNN이 지닌 장기간 메모리 문제를 해결했고, 많은 분야에서 성공적으로 활용되고 있다. LSTM의 성공 이후, LSTM의 변종들이 많이 연구되고 학계에서 발표되었다. 그 중 주목할 것은 조경현이 2014년에 소개한 Gated Recurrent Unit(GRU)이다.



[그림] GRU개요

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

위의 그림에서 알 수 있듯, GRU의 구조가 LSTM의 구조에 비해 간결하다는 것을 확인할 수 있다. 자세히 보면 LSRM과 마찬가지로 gate를 이용하여 정보의 양을 조절하는 것은 동일하지만 gate의 제어방식에 약간의 차이를 보인다.

수식으로 풀어보면, 다음과 같다.

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

Reset gate r_t 는 0~1사이의 값을 가지며, 값이 0에 가까우면 과거의 정보를 모두 잊는다는 것이고, 1이면 과거의 정보를 모두 기억한다. 마지막 줄의 식은 다음 상태로 업데이트를 통제한다. h_{t-1} 는 과거 정보, \tilde{h}_t 는 현재 정보이며 이를 얼마나 조합할지 결정하는 것은 update gate인 z_t 이다.

CNN

CNN은 주로 이미지 인식에 응용되는 신경망 모델링 기법이다. 컴퓨터 비전의 난제라고 여겨졌던 것(물체 카테고리 인식 등)들을 놀랄 정도의 성능으로 해결하고 있다. 기존 완전 연결 신경망의 문제점은 작은 크기의 필기체를 인식하는데 필요한 가중치와 편향의 개수가 폭발적으로 증가한다는 것이다. 예컨대, 16x16 크기의 필기체를 인식하기 위해 단일 은닉층에 100개의 뉴런을 가진 완전 연결 신경망을 구상한다고 해보자. 이 네트워크에 필요한 가중치와 편향은 총 28,326개가 된다. 이 신경망에 은닉층을 한층 더 쌓을 경우, 학습 파라미터의 개수는 더더욱 많아지게 된다. 또한 필기체 이미지가 수직/수평 이동, 회전, 변형이 생기게 되면 새로운 학습데이터를 넣어주어야 하는 문제가 생긴다.

이 문제를 해결하기 위해 탄생한 CNN은 완전 연결 신경망과는 달리 인접한 층과 층 사이에 특정한 유닛만이 결합을 갖는 특수한 층을 가진다. 이 특수한 층에서는 합성곱(Convolution)과 풀링(Pooling)이라는 이미지 처리와 관련한 연산이 발생한다.

CNN의 개발은 동물의 뇌에 있는 시각 피질(Visual cortex)에 대한 신경과학적 지식으로부터 힌트를 얻었다. 더 자세히 표현하자면 시각 피질에 대한 신경세포의 감수 영역(Receptive field)의 국소성과 단순 세포(Simple cell) 및 복잡 세포(Complex cell)의 존재에서 영감을 얻었다고 볼 수 있다. 어떤 뉴런은 수평선 이미지에만 흥분하고, 다른 뉴런은 다른 각도의 이미지에 반응한다. 또 어떤 뉴런은 큰 감수영역을 가지고 있어 저수준 패턴이 조합된 더 복잡한 패턴에 반응한다는 것을 알게 되었다. 이를 이용하여 LeNet, AlexNet, VGG, GoogLeNet 등의 대표적 CNN 모델이 만들어졌으며, 이미지 분류에서 인간 이상의 성능을 보이고 있다.

CNN의 기본구조는 다음과 같이 요약할 수 있다. 입력 이미지를 받는 첫번째 합성곱층의 뉴런은 입력 이미지의 모든 픽셀에 연결되는 것이 아니라 합성곱층 뉴런의 감수영역 안에 있는 픽셀에만 연결된다. 하지만 합성곱은 선형 연산을 실행한다. 따라서 더 복잡한 비선형 관계를 학습시키기 위해, 대표적으로 ReLU(Rectified Linear Unit) 같은 비선형 활성화함수를 합성곱층에 연결시켜준다. 이후 풀링층은 연산량과 메모리 사용량 및 파라미터 수를 줄이기 위해 입력 이미지의 부표본(subsample)을 만들어 주는 역할을 한다. 데이터의 성격에 따라 합성곱층 +

활성함수 + 풀링층의 조합으로 다양하게 층을 쌓을 수 있다. 보통 CNN은 출력전에 1층 이상의 완전 연결층을 붙여주는데, 이는 고차원의 특성 맵의 마지막층을 이미지 분류에 매핑할 수 있게 해준다.

상기 서술하였듯, CNN의 가장 중요한 구성 요소는 합성곱층이다. WXW 크기의 흑백 이미지가 있고, 각 픽셀(Pixel)의 인덱스를 (i,j) 라고 하자. 그리고 HXH 크기의 필터(Filter)가 있고, 필터의 픽셀의 인덱스는 (p,q) 이며 $p,q=0,\dots,H-1$ 라고 하자. 여기서 $H<W$ 이다. 엄밀히 합성곱이란 다음과 같이 정의될 수 있다.

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

합성곱층 다음에 오는 풀링층의 계산 내용은 다음과 같이 간략히 설명될 수 있다. 크기 $WXWXK$ 인 입력 이미지의 픽셀 (i,j) 를 중심으로 하는 HXH 크기의 정사각형 영역에 포함되는 픽셀의 집합을 P_{ij} 라고 하자. 적절히 패딩을 적용하면 이미지의 가장자리도 포함한 P_{ij} 를 구축할 수 있다. 이 P_{ij} 내의 픽셀에 대하여 채널 k 마다 독립적으로 H^2 개 있는 픽셀값을 이용, 하나의 픽셀값인 u_{ijk} 를 구한다. 대표적으로 이 픽셀값을 구하는 방법은 맥스(max) 풀링, 평균 풀링 등이 있다.

맥스 풀링은 H^2 개의 픽셀값 중 최대값을 고르는 방법이며, 다음과 같이 계산한다.

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

평균 풀링은 H^2 개의 픽셀값의 평균을 구하여 픽셀값으로 취한다.

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

마지막으로, 맥스 풀링과 평균 풀링을 아우르는 L_p 풀링은 다음과 같다. 풀링의 연산은 입력 이미지의 각 채널마다 병렬로 실행된다. 즉, 풀링층의 출력 채널 수와 입력 이미지의 채널 수가 같다는 말과 일맥상통한다. L_p 풀링은 수식으로 나타내면

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

이고, $p=1$ 이면 평균 풀링, $p=\infty$ 이면 맥스 풀링이 됨을 쉽게 알 수 있다.

가격 예측 분류문제에서 시계열 데이터의 움직임을 패턴화하여 분석해 볼 수 있다. CNN을 이용하여 구성한 모델은 강력한 예측력을 지닌 것으로 판단된다.

Gradient Boosting

그래디언트 부스팅은 상술한 경사하강법(Gradient Descent)와 부스팅(Boosting)의 합성어이다. 부스팅은 상대적으로 단순한 결정나무(또는 학습기)를 성장시켜 순차적(sequential)으로 학습시키고, 이 결과를 바탕으로 또 하나의 결정나무를 성장시키는 방법이다. 경사하강법이라는 이름에서 알 수 있듯이 그래디언트 부스팅은 손실함수로 미분가능한 모든 함수를 쓸 수 있다.

그래디언트 부스팅의 알고리즘은 다음과 같이 나타낼 수 있다.

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

먼저 학습기(h)를 초기화하고, 각 $m=1,2,\dots,M$ 에 대해 다음 1, 2, 3, 4를 반복한다.

$$r_{im} = - \left[\frac{\partial L(y_i, h(x_i))}{\partial h(x_i)} \right]_{h(x)=h_{m-1}(x)}, \quad i = 1, 2, \dots, m$$

1. 유사 잔차(pseudo-residual)를 계산한다.

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

2. 유사 잔차에 대해 기저 학습기 $h_m(x)$ 를 적합한다.
즉, 훈련데이터는 상단같이 나온다.

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

4. 학습기를 갱신한다.

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

3. 다음과 같은 승수 γ_m 를 계산한다.

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

최종 모델은 다음과 같이 선정한다.

가격움직임 분류문제의 경우, 지수 손실함수를 사용한다면, $r_{im} = y_i \exp(-y_i h_{m-1}(x_i)) = y_i w_{im}$ 가 된다. w_{im} 은 정분류가 된 경우엔 e^{-1} , 오분류가 된 경우에는 e^1 이 된다. 즉, 오분류된 데이터에 더 큰 가중치를 부여하고 다음 단계로 진행하게 되는 것이다. 유사 잔차는 일종의 잔차이므로 부스팅의 일반화는 손실함수가 정의된 경우 최대 경사법에 따라 손실함수를 최소화하는 과정이다.

그래디언트 부스팅 방법을 이용한 모델은 대표적으로 XGBoost, LightGBM, Catboost가 있다.

먼저 XGBoost는 병렬처리를 사용하여 학습과 예측이 빠르고, 다양한 커스텀 최적화 옵션을 제공하며, Greedy 알고리즘을 통한 자동 가지치기로 과적합을 막을 수 있고, 다른 알고리즘과 연계성이 좋다.

다음으로 LightGBM의 가장 큰 장점은 XGBoost보다 학습시간이 매우 짧다는데 있다. 또한 메모리 사용량도 상대적으로 적다. XGBoost보다 2년 후에 개발되어서 장점은 계승하고 단점은 보완하는 방식으로 만들어 졌다고 한다. LightGBM의 리프 중심 분할 방식은 트리의 균형을 맞추지 않고, 최대 손실값을 가지는 리프 노드를 지속적으로 분할하면서 트리의 길이가 길어지고 비대칭적인 규칙 트리가 생긴다. 하지만 이는 균형 트리 분할 방식보다 예측 오차 손실을 최소화할 수 있다.

마지막으로 Catboost는 Ordering Principle의 개념을 대입하여 기존의 타겟 유출(target or data leakage)로 인한 예측 밀림현상(prediction shift)과 범주형 변수에 대한 전처리 문제를 해결했다. 전자에 대해 조금 더 상세히 기술하면, leaf value들을 계산하는 과정에서 Ordered Boosting 기법을 활용하였다. 기존의 그래디언트 부스팅은 새로운 나무를 만들 때 현재 모델에서 사용된 데이터를 재사용하기 때문에 과적합에 취약한 문제점이 있었다. 기존의 방법과 다르게 역순으로 leaf value를 먼저 구하고 나무 구조를 고르는 과정을 채택하였다.

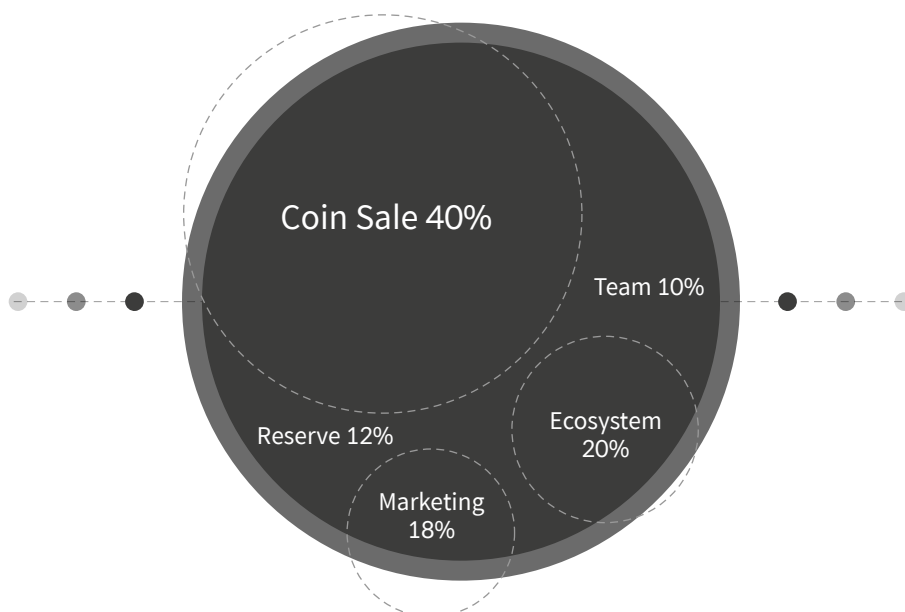
TOKEN ECONOMY

Coin Structure

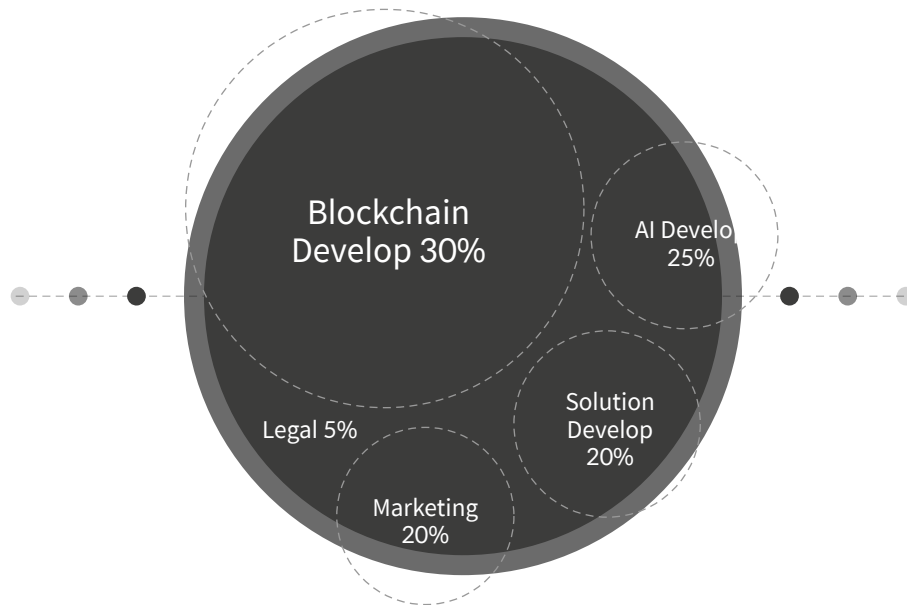
총 발행량	12억개
단위 가격	1AIPE = 약 0.01 USD
코인 이름	AI Prediction & Ecosystem Solution
코인 심볼	AIPE
Standard	ERC-20
Decimal	18
Hard Cap	412,500,000 AIPE
Soft Cap	137,500,000 AIPE

Allocation

AIPE의 발행은 Platform의 개발 및 AI Prediction을 기반으로 한 Ecosystem을 조성하기 위해 이용된다. ICO 및 IEO 참가자들은 전체 발행 량 중 40%에 해당하는 유통량을 지급받게 된다. Coin Sale을 제외한 잔여 물량 중 38%는 초기 Ecosystem 형성에 필요한 AI Prediction Platform 구축 및 Marketing에 이용된다. APIE Reserve에 해당하는 12% 물량은 Platform 추가 개발 및 Platform의 확장됨에 따라 늘어나는 인프라를 구축하기 위해 사용된다. Reserve 물량 12%와 Team 배분 물량 10%는 순차적으로 배분될 예정이다.

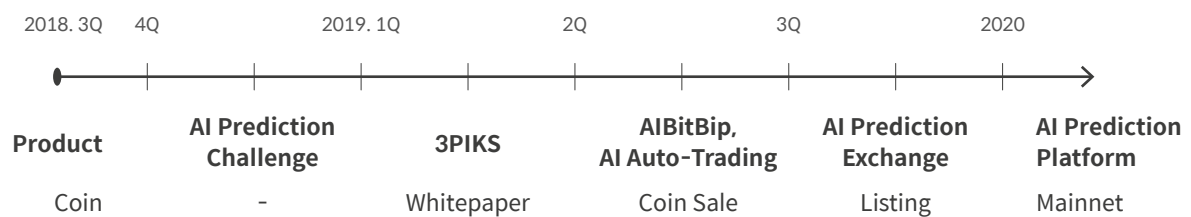


Ecosystem



SERVICE	INFO
Blockchain Develop	블록체인의 연구를 통해 AI Develop Platform을 유지하고 개발된 모델들에 대한 권리보장에 적극 활용 공신력을 갖기 위한 수단으로 AI Prediction Database를 Block에 저장하기 위한 연구를 진행에 지원
AI Develop	AI 연구자들의 교육과 연구를 위한 환경조성 및 예측모델 개발에 적극적인 지원 Machine Learning 기반 Big Data 분석 알고리즘의 지속적 연구
Solution Develop	통합 솔루션을 위한 Product 개발 3PIKS, AIBitBip, AI Auto-Trading Bot, Exchange, Margin Exchange 개발
Marketing	현재 블록체인 산업은 경쟁이 매우 치열한 영역임을 고려 당사의 플랫폼의 빠른 정착을 위해 각종 마케팅 채널을 적극적으로 활용하여 유저를 확보하고 플랫폼을 정착에 이용
Legal	서비스 개발 및 운영상에서 발생하는 각종 저작권 및 관련 법률 처리

ROADMAP



NOTICE & CAUTION

AIPE 프로젝트 백서는 프로젝트의 전반적인 내용과 로드맵을 구체적으로 설명하기 위한 용도로 작성되었습니다. 이 백서는 투자 등을 권유하기 위한 목적으로 작성된 것이 아니며, 이 백서를 읽는 모든 사람들이 이 백서를 참고하여 발생하는 손해, 손실, 채무 등 기타 재무적 피해가 발생하더라도 AIPE는 그에 대한 배상, 보상 기타 책임을 부담하지 않는다는 점에 유의하시기 바랍니다.

이 백서를 읽는 사람이 자신의 의사 결정 등의 행위에 있어서 이 백서를 이용 (이 백서를 참고하거나 이 백서를 근거로 한 경우도 포함하지만 이에 한정되지 아니함) 한 경우에 발생하는 모든 금전적, 채무적 피해에 대해서 AIPE는 그 어떤 배상, 보상 등 기타 책임을 부담하지 않는다는 점에 다시 한번 유의하시기 바랍니다.

AIPE 프로젝트 백서는 ‘작성 당시를 기준(as is)’으로 작성하여 제공 하므로 백서에 포함된 어떠한 내용도 장래 시점까지 정확하거나 변경되지 않는다는 점을 보증하지 않습니다.

AIPE는 이 백서와 관련해서 백서를 읽는 모든 분들에게 어떠한 사항도 진술 및 보장하지 않으며, 그에 대한 법적 책임을 부담하지 않습니다. 예를들어, AIPE는 백서가 적법한 권리에 근거하여 작성되었으며 제 3자의 권리를 침해하지 않는지, 백서가 상업적으로 가치가 있거나 유용한지, 백서가 이 백서를 읽고 있는 사람들이 가지고 있는 특정한 목적의 달성에 적합한지, 백서의 내용에 오류가 없는지 등을 보장하지 않습니다. 책임 면제의 범위는 언급한 예시에만 한정되지 않습니다.

