

퍼지신경망 모형을 이용한 헤지펀드의 생존여부 예측[†]

이광재¹ · 이현준² · 오경주³

¹²³연세대학교 정보산업공학과

접수 2015년 6월 2일, 수정 2015년 7월 1일, 게재확정 2015년 11월 9일

요약

글로벌 금융 위기 발생으로 헤지펀드의 영향력이 증가하면서 헤지펀드의 위험도와 생존여부를 가늠할 새로운 접근법이 필요하게 되었다. 본 연구에서는 헤지펀드의 데이터를 입력값으로 하는 퍼지신경망 모형을 통해 헤지펀드의 생존여부를 예측한다. 헤지펀드의 데이터는 그 변수가 불명확하고 내재적인 불확실성을 가지고 있어 생존 여부의 경계를 설정하는데 어려움이 있다. 따라서 생존 여부를 소속정도로 평가하여 불확실성을 모사할 수 있는 퍼지신경망 모형을 적용하여 예측하고 정확도를 평가한다. 또한 다른 인공지능 방법론들을 이용하여 평가한 결과와 제시한 모형의 성과를 비교하여 그 차이점을 확인한다. 본 연구의 실험결과를 통해 퍼지신경망 모형의 예측력을 확인했으며, 향후 투자자들이 헤지펀드 투자에 대한 의사를 결정하는데 도움을 줄 것으로 기대한다.

주요용어: 생존여부예측, 퍼지신경망모형, 헤지펀드.

1. 서론

세계 금융시장에서 2008년 금융위기를 계기로 헤지펀드의 영향력이 증가함과 동시에 헤지펀드에 대한 규제 및 감독강화에 대한 요구가 늘어나는 등 헤지펀드에 대한 우려 또한 커져가고 있다. 그러나 2005년 9,000여개의 헤지펀드를 시작으로 2007년에는 10,000여개, 그리고 2008년 금융위기 때 성장세가 주춤하다가 매 해 꾸준히 증가하여 2014년 기준 11,000여개의 헤지펀드가 운용되고 있으며 그 규모는 2조 6천억 달러를 기록했고, 금융의 국제화를 고려할 때도 헤지펀드에 대해 부정적으로만 받아들일 수는 없는 상황에 이르렀다.

헤지펀드의 이러한 성장은 헤지펀드가 여타 펀드에 비해 수익률이 높고 포트폴리오 편입 시 위험이 감소하며 수익률 프로파일이 개선된다는 긍정적 효과를 배경으로 한다. 또한 창의적인 금융 인력을 양성할 수 있고 금융업에서의 고용 기회가 확대된다는 점에서 산업 발전에 이바지 한다는 해석도 가능하다. 한편 헤지펀드는 시장 전체의 안정성을 저해하는 시스템리스크 발생과 투자자 보호 및 의결권 행사 등에 있어 논란이 제기되는데, 시스템리스크는 시장의 안정성을 저해할 수 있고 금융시장의 위험이 실물경제에까지 전이 될 수 있으며 일반적인 간접투자기구에 비해 규제가 상당 부분 완화되는 헤지펀드의 경우 금융 감독의 사각지대에 놓일 우려가 있어, 규제회피 수단으로 악용되어 투자자 보호가 제대로 이루어지지 않을 수 있다는 부정적인 면도 존재한다.

[†] 이 논문은 2014년 정부 (미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2014R1A2A1A11052670).

¹ (120-749) 서울특별시 서대문구 신촌동 134번지, 연세대학교 정보산업공학과, 석사과정.

² (120-749) 서울특별시 서대문구 신촌동 134번지, 연세대학교 정보산업공학과, 석사과정.

³ 교신저자: (120-749) 서울특별시 서대문구 신촌동 134번지, 연세대학교 정보산업공학과, 교수.

E-mail: johanoh@yonsei.ac.kr

긍정적 효과와 부정적 효과가 헤지펀드에 내재하고 있기 때문에 상호 배타적인 시각이 존재하나, 금융 시장에서 영향력이 크게 증가한 것은 분명한 사실이다. 따라서 이에 따르는 헤지펀드의 손실과 역할에 대한 논란이 가중되어 왔다. 헤지펀드의 시장리스크에 관한 대표적 사례로 Amaranth Fund (2006)의 경우 천연가스 선물거래로 총 자산인 95억 달러의 65%에 해당하는 60억 달러의 투자손실을 입은 직후 파산한 사례가 있다. 한편 헤지펀드 금융사기의 대표사례로 Bayou Capital은 거래브로커 회사를 지원하기 위해 유동성이 낮은 증권시세를 조작하여 투자자들로부터 5억 달러를 횡령한 사례가 있다.

헤지펀드의 실패 사례에서 헤지펀드의 투자위험과 유동성에 대한 연구가 부족함을 알 수 있다. 특히 인공지능을 이용한 헤지펀드의 실패 연구는 찾아보기 어려운 실정이다. 따라서 본 연구에서는 헤지펀드의 실패를 예측하는 데 인공지능 기법을 적용하여 펀드의 생존과 실패를 예측한다. 헤지펀드 데이터의 경우 비선형적이고 수리적으로 모형화 하기 힘든 특징이 있기 때문에 이를 해결하는 데 유리한 퍼지시스템을 도입하고 퍼지 규칙을 작성하는 데 필요한 지식획득 과정을 자동화 시킬 수 있는 인공지능망 모형을 결합한 퍼지신경망 모형을 이용한다. 이 중 탁월한 함수근사능력을 갖추어 시계열 예측과 시스템 컨트롤 분야에 널리 사용되고 있는 모형은 신경망 기반 퍼지추론시스템 (adaptive network-based fuzzy inference system; ANFIS)이다 (Jang, 1993). 먼저 주식시장 트렌드 예측에서 신경망 기반 퍼지추론 시스템을 적용하여 다른 예측모형 대비 우수한 결과를 보인 연구가 Atsalakis와 Valavanis에 의해 진행되었다 (Atsalakis와 Valavanis, 2009). 따라서 본 실험에서는 ANFIS모형을 사용하여 헤지펀드의 생존 여부를 예측해 본다. 또한 널리 사용되고 있는 인공지능 방법론인 인공지능망, 의사결정나무, 로지스틱 회귀분석을 통해 헤지펀드의 생존여부를 예측해보고 그 결과를 퍼지신경망의 결과와 비교해보고자 한다.

본 논문은 다음과 같이 구성되어 있다. 2절에서는 퍼지신경망 및 비교를 위해 사용된 다른 인공지능 방법론들에 대해 살펴보고, 3절에서는 이 논문에서 제시하는 분석 모형에 대해 단계별로 설명하였으며, 4절에서는 제안한 예측모형의 실증분석 결과를 분석하였다. 마지막으로 결론에서는 본 연구의 기대효과 및 향후 연구에 대해 서술하였다.

2. 연구배경

2.1. 퍼지신경망

ANFIS모형은 기존의 인공지능망과 퍼지시스템을 모두 이용한다. ANFIS는 다수의 If-Then 퍼지 규칙과 규칙에 상응하는 입력값과 출력값들로 구성된다. 또한 ANFIS는 인공지능망의 훈련알고리즘 (training algorithms)과 학습알고리즘 (learning algorithms)을 구성하여 이용한다.

두 개의 입력과 한 개의 출력을 가지는 일차선형 Sugeno퍼지모형을 예로 들면 다음과 같은 규칙으로 구성된다. 여기서 A_1 은 퍼지집합을 나타내며 f_1 은 출력함수를, p_1 , q_1 , r_1 은 출력함수의 모수이다.

$$\text{만약 } x \text{가 } A_1 \text{이고 } y \text{가 } B_1 \text{이면 } f_1 = p_1x + q_1y + r_1$$

ANFIS를 기반으로 하는 분류모형은 다섯 계층의 전방향 연결 신경망 구조로 되어 있고 이는 Figure 2.1에 나타나있다. 각 계층별 구조와 노드함수를 설명하면 다음과 같다.

계층1. 각 노드 i 는 사각형의 노드로 표시되며 다음과 같은 함수로 표현된다.

$$O_{1,i} = \mu_{A_i}(x), \quad i = 1, 2, 3, \quad O_{1,i} = \mu_{B_{i-3}}(y), \quad i = 4, 5, 6$$

여기서 x 와 y 는 노드 i 로의 입력값이며 A_i 와 B_i 는 이 노드함수와 연관된 퍼지집합이다. 즉 $O_{1,i}$ 는 입력 x 와 y 가 각각 퍼지집합 A_i 와 B_i 에 소속되는 정도를 나타내는 소속 함수이다. 일반적으로 $\mu_{A_i}(x)$ 와 $\mu_{B_i}(x)$ 는 종형 (bell-shaped) 함수로 표현되며 최댓값과 최솟값은 각각 1과 0을 갖고 다음과 같은 식으로 표현된다.

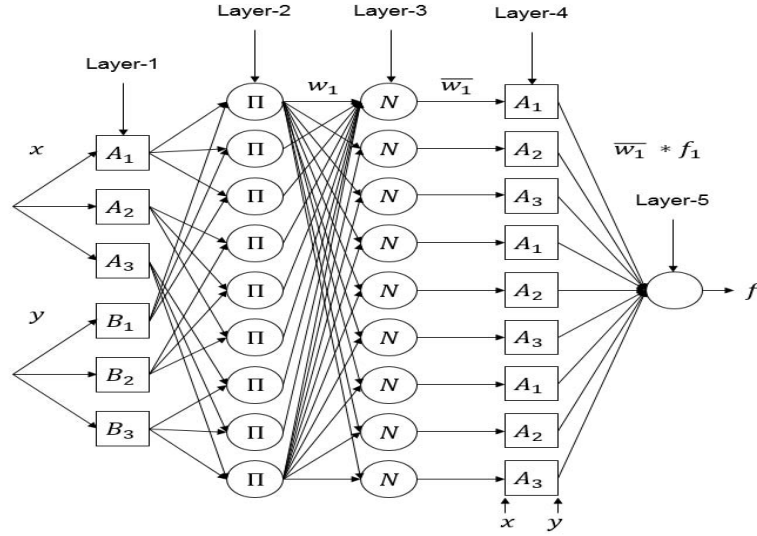


Figure 2.1 ANFIS architecture of two inputs and nine rules

$$\mu_{A_i}(x) = e^{-((x_i - c_i)/a_i)^2}, \quad \mu_{B_{i-3}}(y) = e^{-((y_i - c_i)/a_i)^2}$$

여기서 a_i 와 c_i 는 소속함수의 모수이다. 종형 함수의 각 모수 값이 달라지면 소속함수의 모양도 달라지며 학습과정을 통해 오차를 최소화하는 방향으로 모수 값을 조정하게 된다. 소속 함수의 모수는 퍼지 규칙의 조건부에 나타나기 때문에 조건부 모수로 부른다.

계층2. 원형으로 표시되는 이 계층의 노드들은 이전 계층에서 산출된 소속도를 입력받고 각 소속도를 곱하여 출력한다. 출력값은 퍼지 규칙의 적합도를 나타내며 다음 식으로 표현된다.

$$O_{2,i} = w_i = \mu_{A_i}(x) \cdot \mu_{B_{i-3}}(y), \quad i = 1, 2, 3, \dots, 9$$

계층3. 이 계층의 노드 i 는 i 번째 규칙의 적합도가 전체 규칙들의 적합도 합에 대해 차지하는 비중을 표현하게 되며 다음과 같은 값을 출력한다.

$$O_{3,i} = \bar{w}_i = w_i / (w_1 + w_2 + \dots + w_9), \quad i = 1, 2, 3, \dots, 9$$

계층4. 이 계층의 각 노드에서는 각 규칙 결론부의 출력함수에 표준화된 적합도를 곱하는 연산을 수행한다. 여기서 w_i 는 계층3의 출력값이며 출력함수 모수인 p_i, q_i, r_i 는 결론부 모수라고 부른다. 이 모수 또한 학습과정을 통해 오차를 최소화하도록 결정된다.

$$O_{4,i} = \bar{w}_i \cdot f_i = w_i \cdot (p_i x + q_i y + r_i), \quad i = 1, 2, 3, \dots, 9$$

계층5. 이 계층은 단일노드이며 입력되는 모든 값을 합하여 출력하며 그 값은 연속형으로 표현된다.

$$O_{5,i} = \sum_i \bar{w}_i f_i = \sum_i w_i f_i / \sum_i w_i$$

2.2. 인공신경망

인공신경망은 인간의 두뇌가 문제를 해결하는 방법을 모형화한 것으로 학습, 추론능력이 뛰어나다. 인공신경망은 입력층, 은닉층, 출력층의 삼층구조로 이루어져 있으며 각 층은 뉴런이나 노드라고 불리는 처리단위로 구성되어 있다. 은닉마디 (hidden units)는 인공신경망을 일반적인 통계모형과 구분하는 독특한 구성요소로 인간의 신경세포를 모형화한 것이다. 인공신경망에서 각 은닉마디는 입력변수들의 결합을 받아 목표변수에 전달한다. 이 때 결합에 사용되는 계수 (coefficient)들을 연결강도 (synaptic weights)라고 부르며 활성화함수 (activation function)는 입력값을 변환하고 이를 입력으로 사용하는 다른 마디로 출력하게 된다.

본 실험에서는 다층 전향구조 (multi-layered feedforward) 인공신경망을 이용하였고 대부분의 문제에서 만족할 만한 수준을 보이는 3층 퍼셉트론 구조를 이용한다.

2.3. 의사결정나무

의사결정나무 모형은 의사결정규칙 (decision rule)을 나무구조로 도표화하여 분류 (classification)와 예측 (prediction)을 수행하는 분석 방법으로 분류와 예측의 과정이 나무구조에 의한 추론규칙 (induction rule)에 의해 표현된다. 연구자가 그 과정을 쉽게 이해하고 설명이 가능하며 어떤 변수가 특정상태의 분류에 영향을 많이 주는지 쉽게 파악이 가능하다. 본 실험과 같이 이산형 목표변수를 가질 때는 분리기준을 엔트로피 지수 (entropy index)로 선택하며 S 는 주어진 데이터들의 집합, C_i 는 클래스 값들의 집합, $freq(C_i, S)$ 는 S 에서 클래스 C_i 에 속하는 레코드의 수, 절대값 S 는 주어진 데이터들의 집합의 데이터 개수를 의미하고 다음 식과 같이 표현된다.

$$Entropy(S) = \sum_{i=1}^m p_i \log_2(p_i), \quad p_i = \frac{freq(C_i, S)}{|S|}$$

2.4. 로지스틱 회귀분석

로지스틱 회귀분석은 목표변수가 이항형이거나 순서형일 경우 사용하는 통계적 분석방법이다. 회귀분석의 목표변수가 입력변수들에 의해서 어떻게 설명 또는 예측되는지 알아보기 위해 자료를 적절한 함수 식으로 표현하여 분석하는 방법으로 본 연구와 같이 목표변수가 1과 0인 이항형의 경우 선형회귀모형의 단점을 극복하기 위해 아래와 같은 로짓변환을 고려한다 (Lee와 Chun, 2013).

$$\ln \frac{p(y=1|x)}{1-p(y=1|x)} = \alpha + \beta x$$

단순 로짓모형은 입력변수의 개수가 P개인 경우에도 사용되므로 일반적인 판별분석과 비교하여 로지스틱 판별분석 (logistic discrimination)이라고 부른다.

본 실험에서는 연결함수 (link function)가 $p(x) = e^x / (1 + e^x)$ 를 따르는 로지스틱 모형을 이용한다.

3. 연구방법

본 절에서는 제안한 분석 모형의 구조와 특징에 대해 설명한다. 모형의 1단계에서는 헤지펀드 데이터를 가공하고 모형의 입력값으로 사용할 변수를 선택한다. 2단계에서는 퍼지신경망 모형의 기본 소속함수를 선택하고 학습방법을 설정한다. 3단계에서는 학습 알고리즘을 통한 규칙을 생성하고 모수를 조정한다. 마지막 단계에서는 헤지펀드별 추론값을 산출하여 적중률을 측정하고 다른 방법론들의 결과와 비교한다.

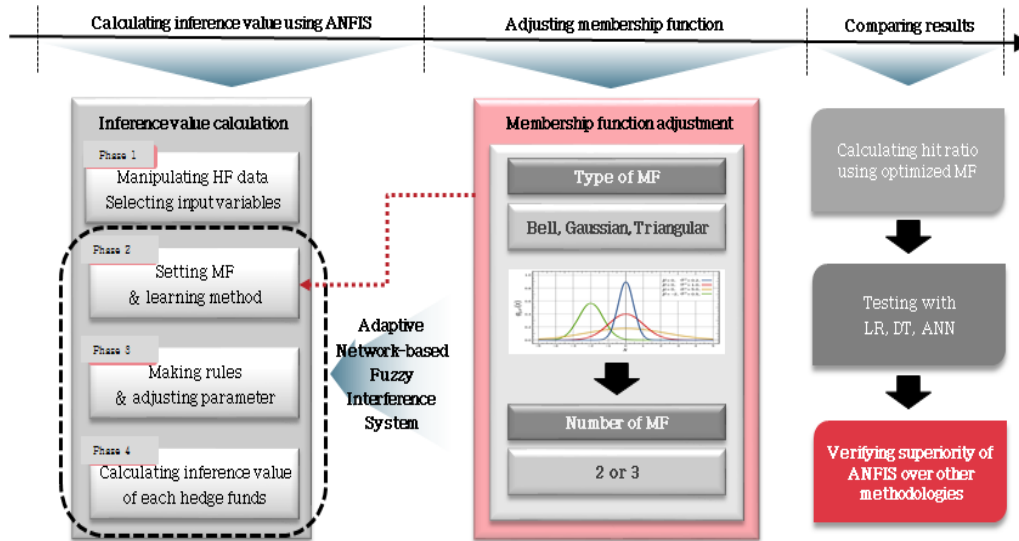


Figure 3.1 The proposed experiment model of this study

1단계. 데이터의 가공 및 입력변수의 선정

본 연구에서 사용한 헤지펀드 연구 데이터베이스는 가공을 거치지 않고 보고된 자료로 신뢰성있는 실험 결과를 얻기 위해서는 데이터에 존재하는 편향을 고려하여 데이터를 가공해야 한다. 헤지펀드 데이터에는 새로운 색인 성분의 역사적 데이터가 포함되어 초기 몇 년의 경우 성과가 매우 좋게 보고될 때가 있다. 따라서 최초의 데이터가 등록된 일자와 헤지펀드가 개시된 시기가 일치하지 않는 경우는 실험 데이터에서 제외한다. 또한 제때 보고가 이루어지지 않아 기록이 없는 구간이 있는 헤지펀드 또한 제외한다. 그리고 실험결과인 적중률의 해석을 용이하게 하기 위해 살아있는 헤지펀드와 운용이 중단된 헤지펀드의 수를 동일하게 실험 데이터를 구성한다.

실험에 사용될 입력변수로는 과거의 연구를 참고하여 헤지펀드의 성과와 생존에 영향을 크게 미치는 변수들을 선정하였다 (Kitabo와 Kim, 2014). 헤지펀드의 특성을 6개의 기준으로 분류하고 중요성이 높다고 알려져 각 기준을 대표할 수 있는 변수들을 선정하였다. 그 결과 총 15개의 헤지펀드 변수 중 8개의 변수를 최종 선별했다. 이는 실험결과에 영향을 미칠 수 있는 불필요 변수를 없애 보다 정확한 결과를 얻을 수 있게 하고, 변수의 수가 증가할수록 규칙의 수와 소속함수의 조합이 기하급수적으로 늘어나는 시스템 특징을 고려할 때 원활하게 실험을 진행할 수 있는 환경을 제공한다.

Table 3.1 Input variables

Classification	Variables
Performance measures	Winning ratio, Average return
Return risk measures	Standard deviation
Fund size measures	Minimum investment
Liquidity	Lockup period
Leverage	Leverage
Fee structure	Management fee, Incentive fee

2단계. 기본 소속함수 선정 및 학습방법 설정

ANFIS의 구성을 위해서는 기본 소속함수의 종류와 그 수를 결정하고 학습방법을 설정해야한다. 소속함수의 형태로는 삼각형 (triangular), 사다리꼴 (trapezoidal), 종형 (bell), 가우시안 (Gaussian), S자형 (sigmoidal) 함수 등이 있으며 이 중 폭넓게 사용되고 있는 삼각형, 종형, 가우시안 함수를 소속함수로 선정한다. 또한 실험 데이터의 특성에 따라 높은 예측도를 보이는 소속함수의 수가 존재할 수 있으므로 각 함수별로 2개와 3개의 소속함수를 생성하여 실험을 진행한다. 학습방법으로는 퍼지 규칙의 해석이 용이하고 입력공간의 차원이 낮을 때 유리한 그리드 분할 (grid partitioning) 방법을 이용한다.

3단계. 학습 알고리즘을 통한 규칙생성 및 모수 조정

구성된 ANFIS에 훈련데이터와 검증데이터를 입력값으로 하여 올바른 예측을 할 수 있는 소속함수의 형태와 퍼지규칙을 생성할 때 까지 학습을 반복한다. 이는 인공신경망의 기능이 동작하는 단계로 훈련횟수 (epoch)는 100회로 하여 오차가 더 이상 발생할 수 없을 만큼의 충분한 수의 학습을 진행하도록 한다. 훈련이 종료되면 소속함수의 모수들이 조정되고 퍼지규칙이 완성되어 실험데이터를 입력하였을 때 헤지펀드의 생존여부를 예측할 수 있는 ANFIS모형이 구성된다.

4단계. 헤지펀드별 추론값과 적중률 산출 및 방법론 별 비교

완성된 ANFIS에 실험데이터를 입력하여 생존한 헤지펀드와 운용이 중단된 헤지펀드를 구분해내는 적중률을 계산한다. 훈련데이터와 검증데이터에는 목표값으로 0의 경우 운용이 중단된 헤지펀드, 1의 경우 생존한 펀드로 구분되어 있으므로 ANFIS의 추론값은 0과 1사이의 연속된 값이 출력된다. 따라서 출력된 값을 0.5를 기준으로 하여 그보다 낮으면 운용이 중단된 펀드, 그보다 높으면 생존한 펀드로 구분한다 (Shim 등, 2012). 인공신경망, 의사결정나무, 로지스틱 회귀분석도 동일한 데이터로 실험을 진행하고 그 결과를 ANFIS의 예측 결과와 비교하여 ANFIS의 성과를 확인한다.

4. 실증분석

이 연구에서 사용된 데이터는 HFR사의 데이터베이스로 1990년 1월부터 2009년 12월까지 보고된 20년간의 자료를 사용하였다. 사용된 데이터베이스는 전 세계의 헤지펀드 매니저들이 직접 보고한 데이터와 HFR사에서 현장 인터뷰를 통해 얻은 7500여개 이상의 헤지펀드들의 정보로 구성되어있다. 데이터 가공을 위해 우선 가공 전의 데이터베이스에서 36개월 미만 펀드 및 데이터베이스 등록일자와 개시일자가 미일치하여 성과에 바이어스가 발생할 수 있는 데이터를 삭제하였다. 그리고 각 집단의 샘플 수를 맞추기 위해 살아있는 펀드와 죽은 펀드 각각 1559개의 데이터를 선택했다. 훈련데이터와 검증데이터, 실험데이터의 세트는 전체 데이터에서 무작위로 추출하여 구성하였다. 비율은 30:30:40으로 훈련데이터, 검증데이터 세트가 각 467개의 데이터, 그리고 실험데이터 세트가 625개의 데이터로 구성되었다. 이러한 방법으로 총 10회에 걸쳐 무작위로 데이터 세트를 구성한 뒤에 10번의 실험을 진행하고 생존 예측 적중률의 평균을 산출하였다. 여기서 적중률은 실험데이터에 포함된 헤지펀드의 수 대비 생존을 올바르게 예측한 경우의 비율을 의미한다.

Table 4.1에서 볼 수 있듯이 ANFIS의 실험 결과 2개의 소속함수로 구성되어 있을 때는 삼각형함수가 적중률이 가장 높고 종형함수, 가우시안함수 순서로 그 적중률이 우수했다. 3개의 소속함수의 경우에는 세 종류의 소속함수 모두 적중률의 증가를 보여 소속함수의 수가 증가할수록 조합되는 규칙의 수가 증가되고 이에 따라 추론값이 더 세분화 된 기준에 의해 결정되었음을 확인할 수 있다. 하지만 삼각형함수나 가우시안함수보다는 종형함수에서 적중률의 증가가 더 두드러지게 나타났는데, 이는 종형함수가 삼각형함수와 비교할 때는 삼각형함수의 선형과 다르게 부드러운 곡선 형태로 이루어져 있기 때문이

며 모수가 2개인 가우시안함수에 비해 3개의 모수를 가짐으로써 가우시안함수와 동일하게 곡선형태의 함수임에도 불구하고 더 민감하게 학습효과에 영향을 받는다는 것을 의미한다.

Table 4.1 Hit ratio of various membership functions

Membership function (MF)	Triangular		Bell		Gaussian	
Number of MF	2	3	2	3	2	3
Hit ratio	0.5747	0.6051	0.5815	0.6148	0.5707	0.5977
	0.5934	0.6074	0.5821	0.6215	0.5771	0.5782
	0.5917	0.5883	0.5926	0.6116	0.5777	0.5977
	0.5864	0.6068	0.5868	0.6253	0.5771	0.5888
	0.5799	0.6003	0.5949	0.6315	0.5742	0.5941
	0.5805	0.5848	0.5780	0.6259	0.5794	0.5823
	0.5875	0.6039	0.5856	0.6022	0.5905	0.5811
	0.5828	0.5872	0.5844	0.6178	0.5841	0.5838
	0.5916	0.5895	0.5803	0.6123	0.5864	0.5829
	0.5875	0.5913	0.5792	0.6185	0.5713	0.5817
	Average	0.5856	0.5965	0.5845	0.6181	0.5789
Standard deviation	0.0057	0.0086	0.0053	0.0080	0.0061	0.0069

가장 우수한 결과를 보인 3개의 종형함수로 구성된 ANFIS와 다른 인공지능 방법론을 비교해 본 결과를 Table 4.2를 통해서 확인해 볼 수 있다. 비교 대상으로 선정한 3개의 방법론 중에서는 로지스틱 회귀 분석이 가장 우수한 적응률을 보였으며 그 다음으로 인공신경망, 의사결정나무 순서로 적응률이 높게 측정되었다. 이전 실험한 3개의 종형함수로 구성된 ANFIS의 적응률은 61.8%로 다른 인공지능 방법론들과 비교했을 때 약 3%에서 크게는 6%까지 향상되었다. 따라서 헤지펀드의 생존여부를 예측하는 데에는 학습을 통해 규칙을 구축하고 이를 통해 예측하는 퍼지신경망 모형이 다른 방법론들에 비해 효과적임을 알 수 있다. 10회에 걸친 실험의 표준편차를 고려하였을 때도 방법론 별 큰 차이가 없어 비교적 안정적인 추론값이 산출되었다고 볼 수 있다.

훈련데이터와 검증데이터, 실험데이터에 대해 그 적응률을 확인해 보면 훈련데이터의 적응률이 높게 측정되었던 방법론 순서로 실험데이터의 결과도 좋게 나왔음을 알 수 있다. 종형함수 3개로 구성된 ANFIS의 경우 실험데이터의 결과도 가장 우수했지만 훈련데이터나 검증데이터의 적응률도 가장 높았다. ANFIS는 규칙을 구성하고 소속함수의 형태를 갖추는 학습데이터와 검증데이터에 대해서도 좋은 결과를 보여주었고 새로운 헤지펀드 데이터에 대해서도 생존여부를 예측하는 능력이 우수함을 알 수 있다.

Table 4.2 Result of comparison with other methodologies

Methodology	ANFIS (Bell-3)	Neural Network	Decision Tree	Logistic Regression
Hit ratio	0.6148	0.5781	0.5490	0.5895
	0.6215	0.5666	0.5602	0.5743
	0.6116	0.5735	0.5580	0.5965
	0.6253	0.5712	0.56519	0.5930
	0.6315	0.5700	0.5569	0.5766
	0.6259	0.5660	0.5490	0.5772
	0.6022	0.5729	0.5479	0.5801
	0.6178	0.5873	0.5518	0.5959
	0.6123	0.5735	0.5636	0.5872
	0.6185	0.5649	0.5658	0.5848
	Average	0.6181	0.5724	0.5564
Standard deviation	0.0080	0.0063	0.0064	0.0078

Figure 4.1를 통해 확인하면 상대적으로 적중률이 낮았던 2개의 소속함수로 구성된 ANFIS와 비교해도 세 종류의 소속함수로 구성된 ANFIS 모두 비교 방법론 중 가장 우수했던 로지스틱 회귀분석의 결과와 크게 차이가 나지 않는 결과를 보여주었다. 따라서 데이터의 특징을 올바르게 파악하고 그 특징에 걸맞게 퍼지신경망을 구성하면 다른 방법론들과 비교할 때 예측문제를 해결하는데 있어 우수한 방법론이 될 수 있을 것이라고 볼 수 있다.

Table 4.3 Hit ratio of training, validation, test sets

Methodology	ANFIS (Bell-3)	Neural Network	Decision Tree	Logistic Regression
Training set	0.6426	0.6081	0.5804	0.6101
Validation set	0.6142	0.5750	0.5710	0.5879
Test set	0.6181	0.5724	0.5564	0.5855

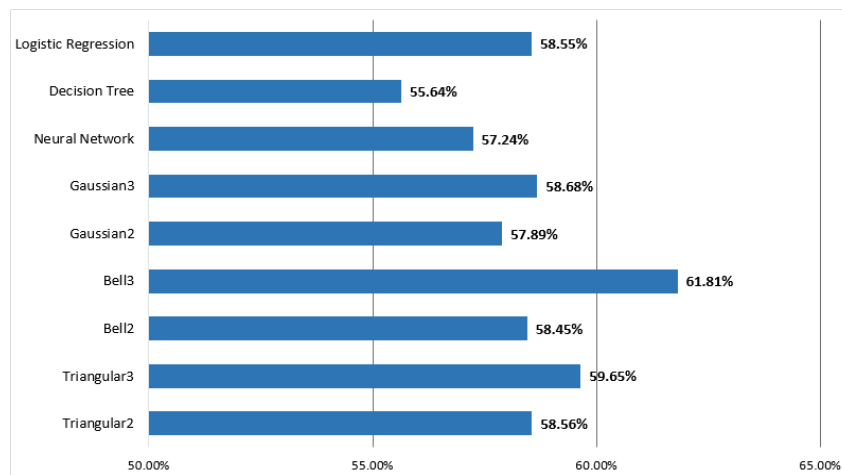


Figure 4.1 Hit ratio of various methodologies

5. 결론

2011년 헤지펀드 국내도입 이후로 현재의 저금리 기조와 적절한 투자처를 찾지 못한 유동성이 헤지펀드로 투자수요를 확대하고 있다. 본 연구는 헤지펀드의 영향력 증가에 따라 헤지펀드의 긍정적 또는 부정적 효과에 대한 논란이 가열되는 상황에서 퍼지신경망 모형을 이용하여 헤지펀드의 생존여부를 예측해보고자 하였다. 퍼지신경망모형 중 예측분야에서 각광받고있는 ANFIS모형을 이용하여 실험을 진행했으며 그 결과를 인공신경망, 의사결정나무, 로지스틱 회귀분석을 통해 얻은 결과와 비교했다. 제 4절에서 보여준 실험결과를 통해 ANFIS모형을 이용한 헤지펀드의 생존여부 예측이 다른 인공지능 방법론 대비 향상된 결과를 얻을 수 있음을 확인했으며 표준편차를 통해 추론값이 안정적으로 산출되었음을 알 수 있었다.

본 연구에서는 ANFIS를 헤지펀드 데이터에 적용하여 그 우수함을 확인했지만 적용범위를 다른 상품이나 시장으로 넓힌다면 투자자들이 투자 의사를 결정하는데 유용하게 이용할 수 있을 것이라고 기대해 본다. 물론 투자 대상에 따라서 고려해야하는 데이터는 그 특성이 다양하여 ANFIS모형을 구성할 때 필요한 소속함수의 종류와 수, 학습방법 등을 새로이 연구하고 실험을 통해 검증해야 할 필요가 있다. 또한 본 연구에서 비교대상으로 제시하였던 방법론 외에 다양하게 발전된 형태의 알고리즘과의 비교해

고 더 우월한 모형을 제시, 연구해 볼 수 있다. 본 실험은 사용된 데이터는 미국을 중심으로 한 해외 헤지펀드 데이터들을 주축으로 진행되었기 때문에 향후 연구에서는 충분히 수집된 국내 헤지펀드의 데이터를 이용한 검증은 통해 국내 시장에서의 예측 능력 또한 평가해 볼 수도 있을 것이다. 다양한 상품에의 적용과 분석을 통해 추후 연구에서는 더 안정적이고 높은 예측력을 보일 수 있는 모형이 나올 수 있기를 기대한다.

References

- Ackermann, C., R. McEnally and D. Ravenscraft. (1999). The performance of hedge funds: Risk, return and incentives. *Journal of Finance*, **54**, 833-874.
- Agarwal, V., N. D. Daniel and N. Y. Naik. (2009). Role of managerial incentives and discretion in hedge fund performance. *Journal of Finance*, **64**, 2221-2256.
- Ammann, M. and P. Moerth. (2005). Impact of fund size on hedge fund performance. *Journal of asset management*, **6**, 219-238.
- Atsalakis, G. S. and K. P. Valavanis. (2009). Forecasting stock market short-term trends using a neuro-fuzzy based methodology. *Expert Systems with Applications*, **36**, 10696-10707.
- Baba, N. and H. Goko. (2009). Survival analysis of hedge funds. *Journal of Financial Research*, **32**, 71-93.
- Babuska, R. and H. Verbruggen. (2003). Neuro-fuzzy methods for nonlinear system identification. *Annual reviews in control*, **27**, 73-85.
- Bae, J. K. (2010). An integrated approach to predict corporate bankruptcy with voting algorithms and neural networks. *Korean business review*, **3**, 79-101
- Baquero, G., J. Horst and M. Verbeek. (2005). Survival, look-ahead bias, and persistence in hedge fund performance. *Journal of Financial and Quantitative Analysis*, **40**, 493-517.
- Bersini, H. and G. Bontempi. (1997). Now comes the time to defuzzify neuro-fuzzy models. *Fuzzy sets and systems*, **90**, 161-169.
- Ding, B. and H. Shawky. (2005). *Hedge fund performance: 1990-2003*, The Annual Financial Management Association Conference, Chicago.
- Han, J. M. (2008). Legislative proposals on the financial investment law for hedge funds. *Business law review*, **22**, 339-381.
- Hedges, R. J. (2003). Size vs performance in the hedge fund industry. *Journal of Financial Transformation*, **10**, 14-17.
- Jang, J. S. R. (1993). ANFIS: Adaptive network-based fuzzy inference systems. *IEEE Transactions on Systems, Man, and Cybernetics*, **23**, 665-685.
- Kitabo, C. A. and Kim, J. T. (2014). Survival analysis of bank loan repayment rate for customers of Hawassa commercial bank of Ethiopia. *Journal of the Korean Data & Information Science Society*, **25**, 1591-1598.
- Lee, D. and Chun, H. (2013). Analysis of factor of life planner's satisfaction after turnover using the cumulative logit model. *Journal of the Korean Data & Information Science Society*, **24**, 1369-1384.
- Lee, H. S. (2011). Evaluation of financial risk of hedge funds and funds of hedge funds, Ph. D. Thesis, Discipline of Finance, Business School, The University of Sydney.
- Melek Acar Boyacioglu and Derya Avci. (2010). An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange. *Expert Systems with Applications*, **37**, 7908-7912.
- Noh, H. J. (2011). *Hedge fund theory and practice*, Seoul, Parkyoungsa.
- Oh, K. J., Kim, T. Y., Jung, K. and Kim, C. (2011). Stock market stability index via linear and neural network autoregressive model. *Journal of the Korean Data & Information Science Society*, **22**, 335-351.
- Shim, K. S., Ahn, J. J. and Oh, K. J. (2012). Multi-currencies portfolio strategy using principal component analysis and logistic regression. *Journal of the Korean Data & Information Science Society*, **23**, 151-159.
- Song, H. S. and Kim, J. K. (2009). Design and evaluation of ANFIS-based classification model. *Journal of intelligence and information systems*, **15**, 151-165.
- Takagi, T. and M. Sugeno. (1983). Derivation of fuzzy control rules from human operator's control actions. In *Proceedings of the IFAC symposium on fuzzy information, knowledge representation and decision analysis*, 55-60.

Using fuzzy-neural network to predict hedge fund survival[†]

Kwang Jae Lee¹ · Hyun Jun Lee² · Kyong Joo Oh³

¹²³Department of Information and Industrial Engineering, Yonsei University

Received 2 June 2015, revised 1 July 2015, accepted 9 November 2015

Abstract

For the effects of the global financial crisis cause hedge funds to have a strong influence on financial markets, it is needed to study new approach method to predict hedge fund survival. This paper proposes to organize fuzzy neural network using hedge fund data as input to predict hedge fund survival. The variables of hedge fund data are ambiguous to analyze and have internal uncertainty and these characteristics make it challenging to predict their survival from the past records. The object of this study is to evaluate the predictability of fuzzy neural network which uses grades of membership to predict survival. The results of this study show that proposed system is effective to predict the hedge funds survival and can be a desirable solution which helps investors to support decision-making.

Keywords: Fuzzy neural network, hedge fund, survival prediction.

[†] This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014R1A2A1A11052670).

¹ Graduate student, Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Korea.

² Graduate student, Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Korea.

³ Corresponding author: Professor, Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Korea. E-mail: johanoh@yonsei.ac.kr